Salvaging Falsified Instrumental Variable Models*

Matthew A. Masten[†] Alexandre Poirier[‡]

February 7, 2021

Abstract

What should researchers do when their baseline model is falsified? We recommend reporting the set of parameters that are consistent with minimally non-falsified models. We call this the *falsification adaptive set* (FAS). This set generalizes the standard baseline estimand to account for possible falsification. Importantly, it does not require the researcher to select or calibrate sensitivity parameters. In the classical linear IV model with multiple instruments, we show that the FAS has a simple closed form expression that only depends on a few 2SLS coefficients. We apply our results to an empirical study of roads and trade. We show how the FAS complements traditional overidentification tests by summarizing the variation in estimates obtained from alternative non-falsified models.

JEL classification: C14; C18; C21; C26; C51

Keywords: Instrumental Variables, Nonparametric Identification, Partial Identification, Sensitivity Analysis

*First arXiv draft: Dec 30, 2018. This paper was presented at Auburn University, UC San Diego, Texas A&M, Duke, Columbia, Cornell, the Yale MacMillan-CSAP workshop, the University of Mannheim, Penn State, the 2018 Incomplete Models conference at Northwestern University, the 2018 and 2019 Southern Economic Association Meetings, the 2019 IAAE conference, the 2019 Georgetown Center for Economic Research Mini-Conference on Non-Standard Methods in Econometrics, the 2019 CeMMAP UCL/Vanderbilt Conference on Advances in Econometrics, the 29th Annual Meeting of the Midwest Econometrics Group, the 2019 Greater New York Area Econometrics Colloquium, the 2020 Winter Meeting of the Econometric Society, and the 2020 World Congress of the Econometric Society. We thank audiences at those seminars and conferences, the referees, as well as Federico Bugni, Tim Christensen, Allan Collard-Wexler, Joachim Freyberger, Chuck Manski, Arnaud Maurel, Francesca Molinari, Adam Rosen, Pedro Sant'Anna, and Alex Torgovitsky for helpful conversations and comments. We thank Paul Diegert and Peiran Xiao for excellent research assistance. Masten thanks the National Science Foundation for research support under Grant No. 1943138.

[†]Department of Economics, Duke University, matt.masten@duke.edu

[‡]Department of Economics, Georgetown University, alexandre.poirier@georgetown.edu

1 Introduction

Many models used in empirical research are falsifiable, in the sense that there exists a population distribution of the observable data which is inconsistent with the model. With finite samples, researchers often use specification tests to check whether their baseline model is falsified. Abstracting from sampling uncertainty, the population versions of such specification tests have a persistent problem: What should researchers do when their baseline model is falsified?

In this paper, we provide a constructive way for researchers to salvage a falsified baseline model. To do this, we consider continuous relaxations of the baseline assumptions of concern. By sufficiently weakening the assumptions, a falsified baseline model becomes non-falsified. We define the *falsification frontier* as the set of smallest relaxations of the baseline model which are not falsified. Our main recommendation is that researchers report estimates of the identified set for the parameter of interest under the assumption that the true model lies on this frontier. We call this the *falsification adaptive set* (FAS). This set collapses to the baseline identified set or point estimand when the baseline model is not falsified. When the baseline model is falsified, this set expands to include all parameter values consistent with the data and a model which is relaxed just enough to make it non-falsified. Hence the FAS generalizes the standard baseline estimand to account for possible falsification. Importantly, researchers do not need to select or calibrate sensitivity parameters to compute the falsification adaptive set. We formally define these concepts in section 2.

To illustrate this method, we study the classical constant coefficients linear model with multiple instruments in section 3. We relax instrument exclusion by allowing the instruments to have some direct effect on outcomes. We show that the FAS has a particularly simple closed form expression, depending only on the value of a handful of 2SLS regression coefficients. We then use our results in an empirical study of roads and trade in section 4. We show that the FAS is an informative complement to traditional overidentification tests: The FAS summarizes the range of estimates obtained from alternative models which are not falsified by the data. Thus the FAS reflects the model uncertainty that arises from a falsified baseline model.

Related Literature

Our paper builds on several large literatures. Manski and Pepper (2018) present identified sets under relaxations of two assumptions, which can be used to construct a falsification adaptive set in their model; see their table 2. Ramsahai (2012) studies a heterogeneous treatment effect IV model with continuous relaxations of instrument exogeneity and informally notes that the model is not falsified if exogeneity is sufficiently relaxed. Machado, Shaikh, and Vytlacil (2019) study a heterogeneous treatment effects IV model and formally define what we call a falsification point. More recently, Andrews and Kwon (2019) introduce a scalar slack parameter to define minimally non-falsified moment inequality models, which do not nest our results.

Our technical results build on a large literature on sensitivity analysis in linear IV models, including Fisher (1961), Angrist and Krueger (1994), Altonji, Elder, and Taber (2005), Small (2007), Conley, Hansen, and Rossi (2012), Ashley (2009), Kraay (2012), Ashley and Parmeter (2015), and van Kippersluis and Rietveld (2017, 2018). There is also a large literature on falsification and sensitivity analysis in heterogeneous effect IV models. See Flores and Chen (2018) and Swanson, Hernán, Miller, Robins, and Richardson (2018) for excellent surveys.

Several recent papers use local asymptotics to study sensitivity to misspecification. For example, see Andrews, Gentzkow, and Shapiro (2017), Bonhomme and Weidner (2018), and Armstrong and Kolesár (2019). This approach assumes the baseline model is approximately correct, in the sense that the magnitude of model misspecification is similar to the magnitude of sampling uncertainty. We focus on clearly falsified models, where it is known that the model is not approximately correct. Hence we use a global approach, which does not rely on linking the size of model misspecification to the size of sampling uncertainty. We compare this local misspecification approach with ours in more detail at the end of section 4.

2 Salvaging Falsified Models

In this section we consider a general falsifiable model. We use this model to precisely define the falsification frontier and falsification adaptive set. In section 3 we illustrate these general concepts in the classical linear instrumental variable model.

2.1 Measuring the Extent of Falsification

Let W be a vector of observed random variables. Let $\mathcal{F}$ denote the set of all possible cdfs for W. A *model* is a set of underlying parameters which generate the observed distribution F_W and restrictions on those parameters. These parameters could be infinite dimensional. This definition of a model suffices for our purposes. See section 2 of Matzkin (2007), for example, for a more formal definition. A given model is *falsifiable* if there are some distributions F_W which could not have been generated by the model. When the data follows one of these population distributions, we say the model is falsified (equivalently, refuted). Let $\mathcal{F}_{\text{f}}$ denote the set of cdfs F_W which falsify the model. Let $\mathcal{F}_{\text{nf}}$ denote the set of cdfs F_W which do not falsify the model.

Suppose we begin with a falsifiable baseline model. Suppose this model has L assumptions which we think might be false. For each assumption $\ell \in \{1, \ldots, L\}$, we define a class of assumptions indexed by a parameter δ_ℓ such that the assumption is imposed for $\delta_\ell = 0$, the assumption is not imposed for δ_ℓ equal to its maximum feasible value $\delta_\ell^{\max}$, and the assumption is partially imposed for $\delta_\ell \in (0, \delta_\ell^{\max})$. Two common values of $\delta_\ell^{\max}$ are 1 and $+\infty$. These assumptions must be nested in the sense that for $\delta'_\ell \geq \delta_\ell$, assumption δ'_ℓ is weaker than assumption δ_ℓ.

Consider the model which imposes assumptions $\delta = (\delta_1, \ldots, \delta_L)$. Let $\mathcal{F}_{\text{nf}}(\delta)$ denote the set of joint distributions of the data which are not falsified by this model. In particular, $\mathcal{F}_{\text{nf}}(0_L)$ denotes the set of joint distributions of the data which are not falsified by the baseline model. Since we assumed the baseline model is falsifiable, $\mathcal{F}_{\text{nf}}(0_L)$ is a strict subset of $\mathcal{F}$. Suppose further that the model which does not impose any of the L assumptions is not falsifiable.

Recall that F_W denotes the observed distribution of the data. Suppose $F_W \notin \mathcal{F}_{\text{nf}}(0_L)$, so that the baseline model is falsified. Partition $\mathcal{D} = [0, \delta_1^{\max}] \times \cdots \times [0, \delta_L^{\max}]$ into two sets:

$$\mathcal{D}_{\text{f}} = \{\delta \in \mathcal{D} : F_W \notin \mathcal{F}_{\text{nf}}(\delta)\} \qquad \text{and} \qquad \mathcal{D}_{\text{nf}} = \{\delta \in \mathcal{D} : F_W \in \mathcal{F}_{\text{nf}}(\delta)\}.$$

$\mathcal{D}_{\text{f}}$ is the set of all assumptions which are falsified. $\mathcal{D}_{\text{nf}}$ is the set of all assumptions which are not falsified. For simplicity assume $\mathcal{D}_{\text{nf}}$ is closed, which holds in our section 3 analysis.

Definition 1. The *falsification frontier* is the set

$$\text{FF} = \{\delta \in \mathcal{D} : \delta \in \mathcal{D}_{\text{nf}} \text{ and for any other } \delta' < \delta, \text{ we have } \delta' \in \mathcal{D}_{\text{f}} \ \}$$

where $\delta' < \delta$ means that $\delta'_\ell \leq \delta_\ell$ for all $\ell \in \{1, \ldots, L\}$ and $\delta'_m < \delta_m$ for some $m \in \{1, \ldots, L\}$.

That is, the falsification frontier is the set of assumptions which are not falsified, but if strengthened in any component, leads to a falsified model. When $L = 1$, the falsification frontier is a singleton called the *falsification point*: For all δ below that point, the model is falsified while for all δ above that point the model is not falsified.

2.2 The Falsification Adaptive Set

Let $\Theta_I(\delta)$ denote the identified set for a parameter of interest $\theta \in \Theta$, given the model which imposes the assumptions δ. When $\delta \in \mathcal{D}_{\text{f}}$, δ is below the falsification frontier. In this case, the identified set $\Theta_I(\delta)$ is empty. When $\delta \in \mathcal{D}_{\text{nf}}$, δ is on or above the falsification frontier. In this case, the identified set $\Theta_I(\delta)$ is nonempty.

Definition 2. Call

$$\bigcup_{\delta \in \text{FF}} \Theta_I(\delta)$$

the *falsification adaptive set*.

The falsification adaptive set is the identified set for the parameter of interest when the true model satisfies one of the assumptions on the falsification frontier. When the baseline model is not falsified, this set collapses to $\Theta_I(0_L)$, the baseline identified set (which may be a singleton). This is what researchers typically report when their baseline model is not falsified. When the baseline model is falsified, however, the falsification adaptive set expands to account for uncertainty about which assumption along the frontier is true. Hence this set generalizes the standard baseline estimand to account for possible falsification.

3 The Classical Linear Model with Multiple Instruments

In this section we illustrate our method in the classical linear instrumental variable model. While many kinds of falsifiable assumptions have been considered in the literature, we focus on the classical case where variation from two or more instruments is used to falsify the model.

3.1 Model and Identification

Let $Y(x, z)$ denote potential outcomes defined for values $(x, z) \in \mathbb{R}^{K+L}$. Assume

$$Y(x, z) = x'\beta + z'\gamma + U \tag{1}$$

where β is an unknown constant K-vector, γ is an unknown constant L-vector, and U is an unobserved random variable. Let X be an observed K-vector of endogenous variables. Throughout we suppose X does not contain a constant. Hence U absorbs any nonzero constant intercept. Let Z be an observed L-vector of potentially invalid instruments. We observe the outcome $Y = Y(X, Z)$. For simplicity we have omitted any additional known exogenous covariates W in equation (1); they can be easily included via partialling out.

Equation (1) imposes homogeneous treatment effects. We also maintain the following relevance and sufficient variation assumptions throughout this section.

Assumption A1 (Relevance). The $L \times K$ matrix $\text{cov}(Z, X)$ has rank K.

Assumption A2 (Sufficient variation). The $L \times L$ matrix $\text{var}(Z)$ is invertible.

A1 implies the order condition $L \geq K$. When there is just one endogenous variable ($K = 1$), A1 only requires $\text{cov}(X, Z_\ell) \neq 0$ for at least one instrument. Other instruments may have zero correlation. In this case, these other instruments provide additional falsifying power. We discuss this further below. If one instrument is an affine combination of the others, A2 does not hold. In this case, just remove affinely dependent instruments until $\text{var}(Z)$ is invertible.

The classical model imposes two more assumptions:

Assumption A3 (Exogeneity). $\text{cov}(Z_\ell, U) = 0$ for all $\ell \in \{1, \ldots, L\}$.

Assumption A4 (Exclusion). $\gamma_\ell = 0$ for all $\ell \in \{1, \ldots, L\}$.

A1–A4 imply that the coefficient vector β is point identified and equals the two stage least squares (2SLS) estimand. Furthermore, these assumptions imply well known overidentifying conditions. The following proposition gives these conditions when there is just a single endogenous variable.

Proposition 1. Suppose $K = 1$. Suppose the joint distribution of (Y, X, Z) is known and satisfies A1 and A2. Then the model (1) with A3 and A4 is not falsified if and only if

$$\text{cov}(Y, Z_m)\,\text{cov}(X, Z_\ell) = \text{cov}(Y, Z_\ell)\,\text{cov}(X, Z_m) \tag{2}$$

for all m and ℓ in $\{1, \ldots, L\}$.

When all instruments are relevant, so that $\text{cov}(X, Z_\ell) \neq 0$ for all $\ell \in \{1, \ldots, L\}$, equation (2) can be written as

$$\frac{\text{cov}(Y, Z_m)}{\text{cov}(X, Z_m)} = \frac{\text{cov}(Y, Z_\ell)}{\text{cov}(X, Z_\ell)}.$$

That is, the linear IV estimand must be the same for all instruments Z_ℓ. This result is the basis for the classical test of overidentifying restrictions (Anderson and Rubin 1949, Sargan 1958, Hansen 1982). Suppose the distribution of (Y, X, Z) is such that the model is falsified. This happens when at least one of our model assumptions fails: (a) homogeneous treatment effects, (b) linearity in X, (c) instrument exogeneity, or (d) instrument exclusion.

Here we maintain the homogeneous treatment effects assumption. We consider models with heterogeneous treatment effects in our working paper Masten and Poirier (2020). We also maintain linearity of potential outcomes in x, which could include known functions of covariates like quadratic terms. In principle our analysis can be extended to allow for relaxations of this functional form assumption, but we leave this to future work.

We thus focus on failure of (c) instrument exogeneity or (d) instrument exclusion as reasons for falsifying the baseline model. These are two different substantive assumptions. Mathematically, however, the same technical analysis can be used to relax both assumptions. For simplicity, here we formally maintain the exogeneity assumption A3 and focus on failure of the exclusion assumption A4.

In general, it is difficult to define a meaningful and tractable class of relaxations of one's baseline assumptions. In the linear model, however, there is a natural way to relax the exclusion restriction. Specifically, we use the following class of assumptions.

Assumption A4′ (Partial exclusion). There are known constants $\delta_\ell \geq 0$ such that $|\gamma_\ell| \leq \delta_\ell$ for all $\ell \in \{1, \ldots, L\}$.

A4′ bounds the magnitude of the direct effect of each instrument on the outcome by known constants. This kind of relaxation of the baseline instrumental variable assumptions was previously considered by Small (2007) and Conley et al. (2012); also see Angrist and Krueger (1994) and Bound, Jaeger, and Baker (1995). Although the instruments may have a direct causal effect on outcomes, the model may nonetheless continue to be falsified for sufficiently small values of the components in δ. For sufficiently large values, however, the model will not be falsified. To characterize the falsification frontier, we begin by deriving the identified set for β as a function of δ.

Theorem 1. Suppose A1–A3 and A4′ hold. Suppose the joint distribution of (Y, X, Z) is known. Then

$$\mathcal{B}(\delta) = \left\{ b \in \mathbb{R}^K : -\delta \leq \text{var}(Z)^{-1}(\text{cov}(Z, Y) - \text{cov}(Z, X)b) \leq \delta \right\} \tag{3}$$

is the identified set for β. Here the inequalities are component-wise. The model is falsified if and only if this set is empty.

The identified set $\mathcal{B}(\delta)$ depends on the data via two terms:

$$\underset{(L\times 1)}{\psi} \equiv \text{var}(Z)^{-1}\,\text{cov}(Z, Y) \qquad \text{and} \qquad \underset{(L\times K)}{\Pi} \equiv \text{var}(Z)^{-1}\,\text{cov}(Z, X).$$

ψ is the reduced form regression of Y on Z. Π is the first stage of X on Z. If we demeaned (Y, X, Z) then we would have $\psi = \mathbb{E}(ZZ')^{-1}\mathbb{E}(ZY)$ and $\Pi = \mathbb{E}(ZZ')^{-1}\mathbb{E}(ZX')$. Theorem 1 shows that the identified set is the intersection of L pairs of parallel half-spaces in $\mathbb{R}^K$. When

$\delta = 0_L$, this identified set becomes the intersection of L hyperplanes in $\mathbb{R}^K$. In this case, β is point identified when $\text{cov}(Z, Y) = \text{cov}(Z, X)b$ for a unique $b \in \mathbb{R}^K$. If $\text{cov}(Z, Y) \neq \text{cov}(Z, X)b$ for all $b \in \mathbb{R}^K$, then the baseline model $\delta = 0_L$ is falsified.

Increasing the components of δ leads to a weakly larger identified set. Furthermore, there always exists a δ with large enough components so that $\mathcal{B}(\delta)$ is nonempty. We characterize the set of such δ below. Before that, we show that the identified set can be written as simple intersection bounds when there is a single endogenous variable.

Corollary 1. Suppose the assumptions of theorem 1 hold. Suppose $K = 1$. Then

$$\mathcal{B}(\delta) = \bigcap_{\ell=1}^{L} B_\ell(\delta_\ell)$$

is the identified set for β, where

$$B_\ell(\delta_\ell) = \begin{cases} \left[\dfrac{\psi_\ell}{\pi_\ell} - \dfrac{\delta_\ell}{|\pi_\ell|}, \dfrac{\psi_\ell}{\pi_\ell} + \dfrac{\delta_\ell}{|\pi_\ell|}\right] & \text{if } \pi_\ell \neq 0 \\ \mathbb{R} & \text{if } \pi_\ell = 0 \text{ and } 0 \in [\psi_\ell - \delta_\ell, \psi_\ell + \delta_\ell] \\ \emptyset & \text{if } \pi_\ell = 0 \text{ and } 0 \notin [\psi_\ell - \delta_\ell, \psi_\ell + \delta_\ell]. \end{cases} \tag{4}$$

Here Π is an L-vector and π_ℓ is its ℓth component.

To interpret this result, first consider an instrument Z_ℓ with a zero first stage coefficient, $\pi_\ell = 0$. If Z_ℓ has a sufficiently strong relationship with the outcome, so that $\psi_\ell \pm \delta_\ell$ does not contain zero, then the model is falsified. Furthermore, in this case falsification can be solely attributed to the assumption that $|\gamma_\ell| \leq \delta_\ell$ for this specific ℓ. This is similar to what is sometimes called the 'zero first stage test' (for example, see Slichter 2014 and the references therein). When this relationship with the outcome is sufficiently small, however, Z_ℓ unsurprisingly has no falsifying or identifying power for β.

Next consider a relevant instrument Z_ℓ, so $\pi_\ell \neq 0$. To interpret corollary 1 in this case, we use the following lemma.

Lemma 1. Suppose $K = 1$. Let $\widetilde{X}_\ell = (Z_1, \ldots, Z_{\ell-1}, X, Z_{\ell+1}, \ldots, Z_L)$. Let e_ℓ be the $L \times 1$ vector of zeros with a one in the ℓth component. Suppose $\pi_\ell \neq 0$. Suppose $\text{cov}(Z, \widetilde{X}_\ell)$ is invertible. Then

$$\frac{\psi_\ell}{\pi_\ell} = e_\ell' \,\text{cov}(Z, \widetilde{X}_\ell)^{-1} \,\text{cov}(Z, Y).$$

This lemma shows that ψ_ℓ/π_ℓ is the population 2SLS coefficient on X using Z_ℓ as the excluded instrument and using the remaining instruments $Z_{-\ell}$ as controls. Thus the identified set $\mathcal{B}(\delta)$ is the intersection of intervals around these 2SLS coefficients using one relevant instrument at a time and controlling for the rest.

Finally, consider the baseline case where $\delta = 0_L$. Corollary 1 implies that $\mathcal{B}(0_L)$ is nonempty if and only if

$$\frac{\psi_m}{\pi_m} = \frac{\psi_\ell}{\pi_\ell}$$

for any $m, \ell \in \{1, \ldots, L\}$ with $\pi_m, \pi_\ell \neq 0$ and $\psi_j = 0$ when $\pi_j = 0$. Moreover, in this case $\mathcal{B}(0_L)$ is a singleton equal to this common value. In this case—when the baseline model is not falsified—we also have

$$\frac{\psi_\ell}{\pi_\ell} = \frac{\text{cov}(Y, Z_\ell)}{\text{cov}(X, Z_\ell)}$$

for all $\ell \in \{1, \ldots, L\}$. That is, ψ_ℓ/π_ℓ equals the population 2SLS coefficient on X using Z_ℓ as an instrument and *not* including $Z_{-\ell}$ as controls. This equality of single instrument 2SLS coefficients with and without controls for the other instruments is an alternative characterization of the classical overidentifying conditions from proposition 1. Note that, when these overidentifying conditions do not hold, it can be shown that the baseline 2SLS estimand is not necessarily in the identified set $\mathcal{B}(\delta)$. Consequently, it will not necessarily be in the falsification adaptive set that we describe below. Instead, as shown via corollary 1 and lemma 1, the identified set depends on the estimands ψ_ℓ/π_ℓ, which use $Z_{-\ell}$ as controls to allow for possible exclusion failures.

3.2 The Falsification Frontier

So far we have characterized the identified set for β given a fixed value of δ, the upper bound on the violation of the exclusion restriction. We now consider the possibility that this identified set is empty when $\delta = 0_L$, so that the baseline model is falsified. Our next result characterizes the falsification frontier, the minimal set of δ's which lead to a non-empty identified set. Here we focus on the single endogenous regressor case. We extend this result to multiple endogenous regressors in the online appendix.

Proposition 2. Suppose A1–A3 hold. Suppose the joint distribution of (Y, X, Z) is known. Suppose $K = 1$. Then the falsification frontier is the set

$$\text{FF} = \left\{ \delta \in \mathbb{R}^L_{\geq 0} : \delta_\ell = |\psi_\ell - b\pi_\ell|, \; \ell = 1, \ldots, L, \; b \in \left[\min_{\ell=1,\ldots,L:\pi_\ell \neq 0} \frac{\psi_\ell}{\pi_\ell}, \max_{\ell=1,\ldots,L:\pi_\ell \neq 0} \frac{\psi_\ell}{\pi_\ell} \right] \right\}. \quad (5)$$

In the proof we show that this set satisfies our definition 1 of the falsification frontier. Specifically: Any $\delta \in \text{FF}$ maps to a non-empty identified set, and strengthening any of the assumptions for a given $\delta \in \text{FF}$ leads to an empty identified set.

3.3 The Falsification Adaptive Set

Next we characterize the falsification adaptive set, which is the identified set for β under the assumption that one of the points on the falsification frontier is true. Here we again focus on the single endogenous regressor case. We generalize to multiple endogenous regressors in the online appendix.

Theorem 2. Suppose A1–A3 hold. Suppose the joint distribution of (Y, X, Z) is known. Suppose $K = 1$. Then

$$\bigcup_{\delta \in \text{FF}} \mathcal{B}(\delta) = \left[\min_{\ell=1,\ldots,L:\pi_\ell \neq 0} \frac{\psi_\ell}{\pi_\ell}, \max_{\ell=1,\ldots,L:\pi_\ell \neq 0} \frac{\psi_\ell}{\pi_\ell} \right] \tag{6}$$

is the falsification adaptive set.

We first sketch the proof of this result and then discuss its implications. It follows from two main steps: First, the identified set $\mathcal{B}(\delta)$ is a singleton for any $\delta \in \text{FF}$ (see lemma 2 in the appendix). Second, each of these singleton sets corresponds to an element in the interval on the right hand side of equation (6) (follows using proposition 2). Thus we obtain the entire interval by taking the union of all these singletons.

Our main recommendation is that researchers report estimates of the falsification adaptive set. Theorem 2 shows that, in the classical linear model we consider here, this set has an exceptionally simple form. Most importantly, no δ's appear on the right hand side of equation (6). This implies that *we can obtain the falsification adaptive set without pre-computing the falsification frontier or selecting any sensitivity parameters*. Furthermore, it is very simple to compute, since it just requires running L different 2SLS regressions.

In this model, we can also immediately see how this set adapts to falsification of the baseline model. When the baseline model is not false, $\psi_m/\pi_m = \psi_\ell/\pi_\ell$ for all $m, \ell \in \{1, \ldots, L\}$ with nonzero π_m and π_ℓ. In this case, the falsification adaptive set collapses to the singleton equal to the common value. This is the same point estimand researchers would usually present when their baseline model is not falsified. As the baseline model becomes more falsified, the values of ψ_ℓ/π_ℓ become more different, and the falsification adaptive set expands. Thus the size of this set reflects the severity of baseline falsification.

3.4 Estimation and Inference

In finite samples, researchers can present sample analog estimates of the falsification adaptive set, along with corresponding confidence sets. Our characterization of the FAS in equation

(6) requires that we first screen for irrelevant instruments. It is not clear how to best do this. We present a first pass approach, but leave a detailed analysis to future work.

Let $\{Y_i, X_i, Z_i\}_{i=1}^n$ be an iid sample from the distribution of (Y, X, Z). Let

$$\mathcal{L}_{\text{rel}} = \{\ell \in \{1, \ldots, L\} : \pi_\ell \neq 0\}$$

be the set of indices corresponding to relevant instruments. Estimate this set by

$$\widehat{\mathcal{L}}_{\text{rel}} = \{\ell \in \{1, \ldots, L\} : F_\ell \geq C_n\}.$$

F_ℓ is the first stage F-statistic when considering Z_ℓ as an instrument and $Z_{-\ell}$ as controls. C_n is a cutoff that diverges as the sample size grows. The assumptions in proposition 3 below ensure that $\widehat{\mathcal{L}}_{\text{rel}}$ is consistent for $\mathcal{L}_{\text{rel}}$.

Let $\widehat{b}_\ell$ be an estimator of ψ_ℓ/π_ℓ, the 2SLS coefficient on X using Z_ℓ as the excluded instrument and $Z_{-\ell}$ as controls. We estimate the falsification adaptive set by

$$\widehat{\text{FAS}} = \left[\min_{\ell \in \widehat{\mathcal{L}}_{\text{rel}}} \widehat{b}_\ell, \ \max_{\ell \in \widehat{\mathcal{L}}_{\text{rel}}} \widehat{b}_\ell\right].$$

We use this estimator in our empirical analysis of section 4. The following result gives conditions under which this estimator is consistent for the FAS.

Proposition 3. For all $\ell \in \{1, \ldots, L\}$ suppose

1. $\widehat{b}_\ell \xrightarrow{p} \psi_\ell/\pi_\ell$ when $\pi_\ell \neq 0$.
2. $F_\ell \xrightarrow{d} \chi^2_1$ when $\pi_\ell = 0$.
3. $F_\ell/n \xrightarrow{p} \kappa_\ell$ when $\pi_\ell \neq 0$, where $\kappa_\ell > 0$ is some positive constant.
4. As $n \to \infty$, $C_n \to \infty$ and $C_n = o(n)$.

Let FAS denote the interval in equation (6). Let d_H denote the Hausdorff distance. Then $d_H(\widehat{\text{FAS}}, \text{FAS}) \xrightarrow{p} 0$.

Assumptions 1–3 hold under standard assumptions on random sampling, existence of moments, and the existence of consistent variance estimators used within F_ℓ. Assumption 4 requires that C_n grows slowly enough to ensure that relevant instruments are kept in $\widehat{\mathcal{L}}_{\text{rel}}$ with probability approaching one. In our empirical analysis we choose $C_n = 10$ as our default cutoff, although we sometimes consider other cutoffs, or a sequence of cutoffs. Inference can be done by using a version of the delta method discussed in Fang and Santos (2019), noting that the min and max are directionally differentiable mappings. We leave a detailed analysis of both the choice of the cutoff and procedures for inference to future work.

4 Empirical Application: Roads and Trade

In this section we apply our results from section 3 to the empirical analysis of roads and trade by Duranton, Morrow, and Turner (2014). They consider a dataset of 66 regions ('cities') in the United States. Their treatment variable is the log number of kilometers of interstate highways within a city in 2007. This variable directly affects the cost of leaving a city, and therefore the cost of exporting from a city: It is easier to export from a city with many kilometers of interstate highways passing through it. Their outcome variable is a measure of how much that city exports. They consider two different ways of measuring exports: Weight (in tons) and value (in dollars). We focus on the weight measure for brevity. They begin by estimating a gravity equation relating the weight of a city's exports to other cities with the highway distance between those cities, both measured in 2007. This equation includes a fixed effect for the exporting city. The estimate of this fixed effect is their main outcome variable. They call this variable the "propensity to export weight." Thus their main goal is to estimate the causal effect of within city highways on the propensity to export weight.

We cannot learn this causal effect by simply regressing the propensity to export weight on within city highways since there is a classic simultaneity problem. We expect that building highways within the city will boost exports. But high export cities may also build more highways to facilitate their existing exports. The authors solve this problem by instrumenting for the number of kilometers of within city highways. They consider three different instruments:

1. *Railroads*: The log number of kilometers of railroads in the city in 1898.
2. *Exploration*: A measure of the quantity of historical exploration routes that passed through the city.
3. *Plan*: The log number of kilometers of highway in the city, according to a planned highway construction map approved by the federal government in 1947. Baum-Snow (2007) had previously used this instrument, and provides a detailed history.

The authors raise concerns about validity of all three instruments. Although they address these concerns with various controls, these controls may still not perfectly fix failures of exogeneity, exclusion, or both. Hence the authors lean on overidentification, stating that

> "Using different instruments, for which threats to validity differ, allows for informative over-identification tests." (page 700)

Table 1: Baseline 2SLS results for Duranton et al. (2014): The effect of highways on export weight. Non-highlighted parts reproduce results from their paper. Highlighted parts are new. Panel A reproduces columns 1–4 of their table 5. It also shows the estimated falsification adaptive set. Panel B uses only two of their instruments, controlling for the other. See text for discussion.

	Dependent variable: Export weight			
	(1)	(2)	(3)	(4)
Panel A. Plan, exploration, and railroads used as instruments				
log highway km	1.13	0.57	0.47	0.39
	(0.14)	(0.16)	(0.14)	(0.12)
log employment		0.52	0.69	0.47
		(0.11)	(0.39)	(0.33)
Market access (export)		-0.45	-0.65	-0.63
		(0.14)	(0.14)	(0.11)
log 1920 population			-0.38	-0.29
			(0.25)	(0.23)
log 1950 population			1.00	0.65
			(0.39)	(0.38)
log 2000 population			-0.74	-0.20
			(0.48)	(0.45)
log % manuf. emp.				0.64
				(0.12)
First stage F stat.	97.5	90.3	80	84.8
Overid. p-value	0.10	0.043	0.15	0.31
$\widehat{\text{FAS}}$	[0.49, 0.86]	[-0.32, 0.28]	[-0.26, 0.31]	[0.18, 0.42]
Panel B. Plan and exploration used as instruments, controlling for railroads				
log highway km	0.79	0.17	0.21	0.23
	(0.24)	(0.20)	(0.15)	(0.14)
log 1898 railroad km	0.33	0.33	0.25	0.16
	(0.15)	(0.12)	(0.12)	(0.10)
First stage F stat.	61.1	65.4	77.8	82.2
Overid. p-value	0.64	0.51	0.48	0.72

Notes: 66 observations per column. All specifications include a constant. Heteroskedasticity robust standard errors in parentheses.

With this motivation, we next present the results.

Results

First consider table 1. In this and all other tables, the non-highlighted parts reproduce results from the original paper. The highlighted parts are new computations which we have added. Panel A reproduces columns 1–4 of table 5 in Duranton et al. (2014). These are their main results. In particular, they are interested in the coefficient on log highway km, the log number of highway kilometers within the city. This coefficient represents their estimate of the causal effect of roads on trade. Here it is estimated by 2SLS, using railroads, exploration,

Table 2: The effect of controlling for unused instruments. Non-highlighted parts reproduce results from their paper. Highlighted parts are new. All columns have employment, market access, and past populations as controls. Columns 4-6 also have manufacturing share of employment as controls. Compare to table 6 of Duranton et al. (2014).

	(1) Plan	(2) Railroads	(3) Exploration	(4) Plan	(5) Railroads	(6) Exploration
Panel A. Without controlling for other instruments						
log highway km	0.49	0.83	0.12	0.38	0.64	0.34
	(0.15)	(0.25)	(0.31)	(0.13)	(0.22)	(0.19)
log % manuf. emp.				0.64	0.60	0.65
				(0.12)	(0.13)	(0.13)
First stage F stat.	141	45.2	14.8	130	40.8	23.8
Panel B. Controlling for other instruments						
log highway km	0.31	4.09	-0.26	0.18	3.65	0.42
	(0.22)	(4.09)	(0.71)	(0.21)	(4.16)	(0.52)
log % manuf. emp.				0.63	0.36	0.61
				(0.12)	(0.38)	(0.12)
log 1898 railroad km	0.21		0.24	0.18		0.16
	(0.12)		(0.11)	(0.11)		(0.11)
log 1528-1850 exploration	-0.053	-0.40		0.025	-0.32	
	(0.077)	(0.36)		(0.065)	(0.40)	
log 1947 highway km		-2.09	0.32		-1.90	-0.13
		(2.50)	(0.46)		(2.48)	(0.36)
First stage F stat.	59.6	1.54	29.1	54.8	1.27	27
$\widehat{\text{FAS}}$ for this specification		[-0.26, 0.31]			[0.18, 0.42]	

Notes: 66 observations per column. All specifications include a constant. Heteroskedasticity robust standard errors in parentheses.

and plan as instruments. At the 10% level, the standard test of overidentifying restrictions passes in the two longest specifications, fails in the second specification, and marginally passes in the first specification. Also note that these specifications do not include all of the additional controls the authors consider; they include those in separate analyses, which we discuss later (our table 3).

We add the estimated falsification adaptive set to these baseline results. This is the last row of panel A. There are two things to notice: First, except for the last specification, none of the 2SLS estimates are within the estimated FAS. This is not surprising since it can be shown that the baseline 2SLS estimand does not need to be inside the FAS. Second, the estimated FAS magnitudes are all generally smaller than the 2SLS point estimates.

To better understand how we computed the estimated FAS, and how to interpret it, next consider table 2. Columns 1–3 include the same baseline controls as column 3 in table 1 while columns 4–6 include the same baseline controls as column 4 in table 1. The only difference is that we no longer use all three variables (plan, railroads, exploration) as instruments.

Instead, in panel A, we use only one of these variables as an instrument and we ignore the other two variables. Panel A reproduces columns 4–6 from table 6 in Duranton et al. (2014). The authors used these results as their main robustness check. They argue that the three estimates 0.38, 0.64, and 0.34 from columns 4–6, panel A, table 2 are consistent with their baseline estimates of 0.47 and 0.39 from columns 3 and 4, panel A, table 1.

However, omitting an invalid instrument can lead to omitted variable bias. In this application we are concerned that some of the instruments may be invalid. Thus the alternative models of interest are those where one of the instruments is valid but the others are not. When computing results in these alternative models, the invalid instruments should be included as controls (see lemma 1). Panel B shows these results. Here we use one instrument while controlling for the other two. For example, in column 1 we use plan as an instrument and control for railroads and exploration.

For brevity, here we only describe the results in columns 4–6. These results use the full baseline specification. Consider column 5, panel B. This result uses railroads as an instrument, controlling for plan and highway. Unlike the uncontrolled result from panel A, railroads is a very weak instrument. Hence we ignore the result using railroads alone, as discussed in section 3.4. Next consider column 4. Here we use plan as the instrument, controlling for the other two. Despite these controls, plan is still a strong instrument. The estimated effect 0.18 is roughly half as large as the estimate from panel A, 0.38. It is also no longer statistically significant at any conventional level. Next consider column 6. Here we use exploration as the instrument, controlling for the other two. Exploration continues to be a strong instrument with these controls. The estimated effect 0.42 in panel B is similar to the effect from panel A, 0.34. It is no longer statistically significant, however.

Putting these coefficient estimates together gives us the estimated FAS, $[0.18, 0.42]$. The endpoints of this set correspond to point estimates from alternative models which maintain validity of only one instrument at a time. The interior of this set corresponds to alternative models which relax validity of all instruments at once, but just enough to avoid falsification.

In panel B of table 2 we found that railroads is a weak instrument when controlling for the other two, and hence it yields the largest point estimates. Given this finding, one may also wonder how removing railroads as an instrument affects the baseline analysis. This is shown in panel B of table 1. All of the coefficients on log highway km are smaller, to the point that they are no longer statistically significant for all but the shortest specification.

Table 3: The effect of controlling for unused instruments, continued. Non-highlighted parts reproduce results from their paper. Highlighted parts are new. This table extends the analysis of table 2 to consider specifications with additional control variables. Columns 1–3 reproduce the results in appendix table 6 of Duranton et al. (2014). Columns 4–6 just add controls for the other instruments.

		Without other IV controls			With other IV controls			
		(1)	(2)	(3)	(4)	(5)	(6)	
Added variable		Plan	Railroads	Exploration	Plan	Railroads	Exploration	$\widehat{\text{FAS}}$
Water	log highway km	0.34	0.66	0.24	0.13	3.96	0.30	[0.13, 0.30]
		(0.16)	(0.26)	(0.29)	(0.21)	(4.56)	(0.49)	
	F stat.	126	25.3	10.9	67.6	1.17	26.2	
Slope		0.39	0.57	0.46	0.20	3.86	0.66	[0.20, 0.66]
		(0.14)	(0.20)	(0.19)	(0.21)	(5.86)	(0.50)	
		133	44.5	22.6	60.3	0.65	25.6	
Census regions		0.32	0.62	0.36	-0.012	3.68	0.40	[-0.012, 0.40]
		(0.14)	(0.12)	(0.20)	(0.24)	(3.24)	(0.64)	
		122	58.3	22.6	39.4	1.59	10.7	
Percent college		0.29	0.56	0.41	0.013	3.64	0.78	[0.013, 0.78]
		(0.13)	(0.23)	(0.18)	(0.19)	(4.13)	(0.49)	
		116	36.6	29.5	47.9	1.16	25.2	
Income per capita		0.35	0.63	0.35	0.079	3.42	0.54	[0.079, 0.54]
		(0.14)	(0.22)	(0.18)	(0.21)	(3.42)	(0.47)	
		123	36.8	26.4	47.8	1.70	24.9	
Percent wholesale		0.41	0.59	0.49	0.22	2.71	0.75	[0.22, 0.75]
		(0.12)	(0.21)	(0.12)	(0.19)	(3.29)	(0.49)	
		136	39.3	23.2	54.4	1.38	25.4	
Traffic		0.42	1.00	0.34	0.23	5.57	0.40	[0.23, 0.40]
		(0.18)	(0.52)	(0.23)	(0.25)	(9.60)	(0.51)	
		79	13.4	44.6	44.3	0.43	26	
All		0.39	0.73	0.58	0.18	2.43	0.69	[0.18, 0.69]
		(0.18)	(0.35)	(0.28)	(0.26)	(3.01)	(0.67)	
		52.6	19.9	15.5	32.1	0.88	6.53	

Notes: 66 observations per column. All specifications include a constant. Heteroskedasticity robust standard errors in parentheses.

Moreover, the standard overidentification tests now all easily pass. (Note that these tests are only comparing results using plan and exploration as instruments.) However, the coefficients on railroads are statistically significant for all but the fourth column. This suggests that the full baseline model using all three instruments could be rejected, and also explains the source of the relatively small overidentification test p-values in panel A.

Thus far we have focused on the baseline results, which do not include all of the possible control variables that the authors discuss. Table 3 shows results with their additional controls. We begin with the full set of baseline control variables, as used in column 4 of table 1. We then add just one control. Each row corresponds to a different control. The last row shows the results that add all controls at once. Unlike the main baseline result, column 4 of table 1, here we only use one instrument at a time. Columns 1–3 use a single instrument, without controlling for the other two. These results reproduce appendix table 6 of Duranton et al. (2014). Based on these results, the authors argue that

> "None of our main results is affected by these controls, even when we use our instruments individually." (page 708)

They also argue that using one instrument at a time is an "even more demanding exercise" than examining the effect of additional controls when using all three instruments (not shown here; see their appendix table 5). As we've discussed, however, omitting the invalid instruments may cause omitted variable bias. So in columns 4–6, we replicate columns 1–3, except now controlling for the other two instruments.

There are three main differences between the results with the instrument controls and those without. First, the railroads instrument is again very weak, leading to large coefficients. This informs our understanding of the results from columns 1–3, since there we observed that the coefficients in column 2 are always larger than those in columns 1 and 3, and are often substantially larger. Second, none of the results are statistically significant at conventional levels. Finally, the coefficients using plan as the instrument all become smaller once the other instruments are controlled for (column 4 versus column 1), while the coefficients using exploration as the instrument all become larger once the other instruments are controlled for (column 6 versus column 3). Thus, ignoring the results using railroads, the overall range of point estimates is larger. This is reflected in the estimated falsification adaptive sets, which are presented in the final column.

Overall, there are two main conclusions from our analysis: First, the evidence suggests that the railroads instrument is the most questionable, and should be used only as a control. Thus the estimates in panel B of table 1 are arguably the most appropriate baseline results. Second, there is substantially more uncertainty in the magnitude of the causal effect of roads on trade than suggested by the original results of Duranton et al. (2014). This is reflected in the various estimated falsification adaptive sets we present. In particular, the estimated FAS for the longest specification is $[0.18, 0.69]$; see table 3. Moreover, accounting for sampling uncertainty would only increase this range. That said, these results do not change the overall qualitative conclusions of the paper: All points in the estimated FAS for the longest specification are still positive, suggesting that the number of within city highways appears to positively affect propensity to export weight.

Comparison with the Andrews et al. (2017) approach

In this subsection we compare our approach with that of Andrews et al. (2017). They study general moment equality models, while we focus on the linear IV model. For the linear IV model, in their example 4 they study the sensitivity of the 2SLS estimator to violations of exclusion or exogeneity of the same magnitude as the sampling uncertainty (proportional to $1/\sqrt{n}$). Under such data generating processes, and as in Conley et al. (2012), they show that the 2SLS estimator is consistent, but asymptotically biased. The asymptotic bias has the form $A\gamma$ where γ is a vector of sensitivity parameters and A is a matrix that is point identified from the data. Andrews et al. (2017) recommend that authors report estimates of A, which allows readers to select a γ and compute their own local asymptotic bias correction for the 2SLS estimator.

Our empirical application has a single endogenous variable, log highway km. We are primarily interested in its coefficient. For a given choice of γ, it can be shown that the local asymptotic bias in the 2SLS estimate of this coefficient is $a_1\gamma_1 + \cdots + a_L\gamma_L$ where

$$a_\ell = \frac{\text{cov}(Z_\ell, X_{\text{pred}}^{\perp W})}{\text{var}(X_{\text{pred}}^{\perp W})}. \tag{7}$$

Here W is the vector of the included exogenous control variables, X_{pred} is the predicted value of the endogenous variable from the first stage regression of X on $(1, Z, W)$, and $X_{\text{pred}}^{\perp W}$ is the residual from the linear regression of X_{pred} on $(1, W)$. Thus the relevant elements a_ℓ of the matrix A are simply the coefficients on X_{pred} in a linear regression of Z_ℓ on $(1, X_{\text{pred}}, W)$ for each $\ell = 1, \ldots, L$. In addition to reporting the elements of A, Andrews et al. (2017) also recommend comparing their magnitudes. As they note, this can be difficult because the units of this matrix depend on the units of the moments themselves. In our case, this means that the instrument ℓ with the largest value of $|a_\ell|$ is not necessarily the most important for the local asymptotic bias—it depends on the units of each Z_ℓ. To address this, we standardize the instruments as described below.

Table 4 shows the empirical results. Here we focus on the four main baseline specifications, as reported in table 1. Panel A reports estimates of $a_\ell/\text{stddev}(Z_\ell)$ for each of the three instruments. We see that plan has the largest value of the three instruments. Given equation (7) above, this means that plan has the largest correlation with the predicted treatment, after controlling for covariates. This is consistent with what we reported in table 2, where plan had the largest first stage F-statistic conditional on the other two instruments. Note, however, that railroads and exploration both have roughly the same values in panel

B. This suggests that they are equally important for the asymptotic bias of the 2SLS estimator. In contrast, our analysis in table 2 suggested that, unlike exploration, railroads is a conditionally very weak instrument, and should possibly not be relied on.

In panel B we consider two possible choices of the local sensitivity parameters: $\gamma_\ell = \pm 1/\text{stddev}(Z_\ell)$. These choices can be interpreted as saying that if Z_ℓ increases by one standard deviation, then the direct effect of Z_ℓ on outcomes is $\pm 1/\sqrt{n}$. For these choices, the magnitude of the estimated asymptotic bias is quite large, leading to bias corrected 2SLS estimators which are both positive and negative for the main specification (column 4). This sensitivity analysis thus suggests that even the main qualitative results of Duranton et al. (2014) are not robust to violations of exclusion of that magnitude. In contrast, the estimated FAS for the main specification (column 4 of table 1) still contains only positive numbers, suggesting that the qualitative results of Duranton et al. (2014) *are* robust to exclusion violations that are sufficiently large to explain falsification of the baseline model.

Table 4: Analysis of the local sensitivity of the 2SLS estimator for the four baseline specifications in table 1.

	Dependent variable: Export weight			
	(1)	(2)	(3)	(4)
Panel A. Estimates of $a_\ell/\text{stddev}(Z_\ell)$, the standardized elements of the matrix A				
Plan	1.80	2.36	2.46	2.46
Railroads	1.44	1.73	1.49	1.45
Exploration	1.09	1.48	1.55	1.41
Panel B. Point estimates of the coefficient on log highway km				
Baseline 2SLS	1.13	0.57	0.47	0.39
Bias ($\gamma_\ell = 1/\text{stddev}(Z_\ell)$)	0.53	0.69	0.68	0.69
Bias corrected 2SLS ($\gamma_\ell = 1/\text{stddev}(Z_\ell)$)	0.60	-0.11	-0.21	-0.30
Bias corrected 2SLS ($\gamma_\ell = -1/\text{stddev}(Z_\ell)$)	1.67	1.26	1.14	1.08

Notes: $n = 66$ observations per column. $1/\sqrt{n} = 0.12$. In panel B, Bias is an estimate of $(a_1\gamma_1 + \cdots + a_L\gamma_L)/\sqrt{n}$.

Overall, this discussion highlights two key practical differences between our analysis and that of Andrews et al. (2017). First, their approach is estimator specific: Different estimators of the same parameter can lead to different conclusions about sensitivity. Thus our comparisons of the relative sensitivity of exclusion violations for each of the three instruments above could change if we used an estimator other than 2SLS. In contrast, our approach is a population level identification analysis that does not depend on a specific choice of baseline estimator. Second, their approach ultimately requires researchers—either authors or readers—to choose the sensitivity parameter γ. In contrast, our approach can be thought of as leveraging falsification of the baseline model to automatically calibrate the parameter γ, by considering the minimal relaxations that make the instruments consistent with each other.

5 Conclusion

In this paper we suggested a constructive answer to the question "What should researchers do when their baseline model is falsified?" We recommend reporting estimates of the set of parameters that are consistent with minimally non-falsified models. We call this the *falsification adaptive set* (FAS) because it generalizes the standard baseline estimand to account for possible falsification. We illustrated this recommendation in the classical linear instrumental variable model with multiple instruments. We showed that the FAS has a particularly simple closed form expression, depending only on the value of a handful of 2SLS regression coefficients. Finally, we showed how to use our results in practice. There we discussed the importance of controlling for the possibly invalid instruments when considering alternative models. Overall, we showed that the FAS is an informative complement to traditional overidentification test p-values: It directly summarizes the range of estimates corresponding to non-falsified alternative models.

References

Altonji, J. G., T. E. Elder, and C. R. Taber (2005): "An evaluation of instrumental variable strategies for estimating the effects of catholic schooling," *Journal of Human Resources*, 40, 791–821.

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

Andrews, D. W. and S. Kwon (2019): "Inference in moment inequality models that is robust to spurious precision under model misspecification," *Cowles Foundation Discussion Paper No. 2184*.

Andrews, I., M. Gentzkow, and J. M. Shapiro (2017): "Measuring the sensitivity of parameter estimates to estimation moments," *The Quarterly Journal of Economics*, 132, 1553–1592.

Angrist, J. and A. B. Krueger (1994): "Why do World War II veterans earn more than nonveterans?" *Journal of Labor Economics*, 12, 74–97.

Armstrong, T. and M. Kolesár (2019): "Sensitivity analysis using approximate moment condition models," *arXiv preprint arXiv:1808.07387*.

Ashley, R. (2009): "Assessing the credibility of instrumental variables inference with imperfect instruments via sensitivity analysis," *Journal of Applied Econometrics*, 24, 325–337.

Ashley, R. A. and C. F. Parmeter (2015): "Sensitivity analysis for inference in 2SLS/GMM estimation with possibly flawed instruments," *Empirical Economics*, 49, 1153–1171.

Baum-Snow, N. (2007): "Did highways cause suburbanization?" *The Quarterly Journal of Economics*, 122, 775–805.

Bonhomme, S. and M. Weidner (2018): "Minimizing sensitivity to model misspecification," *arXiv preprint arXiv:1807.02161*.

Bound, J., D. A. Jaeger, and R. M. Baker (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., C. B. Hansen, and P. E. Rossi (2012): "Plausibly exogenous," *The Review of Economics and Statistics*, 94, 260–272.

Duranton, G., P. M. Morrow, and M. A. Turner (2014): "Roads and trade: Evidence from the US," *The Review of Economic Studies*, 81, 681–724.

Fang, Z. and A. Santos (2019): "Inference on directionally differentiable functions," *The Review of Economic Studies*, 86, 377–412.

Fisher, F. M. (1961): "On the cost of approximate specification in simultaneous equation estimation," *Econometrica*, 29, 139–170.

Flores, C. and X. Chen (2018): *Average Treatment Effect Bounds with an Instrumental Variable: Theory and Practice*, Springer.

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

Kraay, A. (2012): "Instrumental variables regressions with uncertain exclusion restrictions: A Bayesian approach," *Journal of Applied Econometrics*, 27, 108–128.

Machado, C., A. Shaikh, and E. Vytlacil (2019): "Instrumental variables and the sign of the average treatment effect," *Journal of Econometrics*, 212, 522–555.

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

Masten, M. A. and A. Poirier (2020): "Salvaging falsified instrumental variable models," *arXiv preprint arXiv:1812.11598v3*.

Matzkin, R. L. (2007): "Nonparametric identification," *Handbook of Econometrics*, 6, 5307–5368.

Ramsahai, R. R. (2012): "Causal bounds and observable constraints for non-deterministic models," *Journal of Machine Learning Research*, 13, 829–848.

Sargan, J. D. (1958): "The estimation of economic relationships using instrumental variables," *Econometrica*, 26, 393–415.

Slichter, D. (2014): "Testing instrument validity and identification with invalid instruments," *Working paper*.

Small, D. S. (2007): "Sensitivity analysis for instrumental variables regression with overidentifying restrictions," *Journal of the American Statistical Association*, 102, 1049–1058.

Swanson, S. A., M. A. Hernán, M. Miller, J. M. Robins, and T. Richardson (2018): "Partial identification of the average treatment effect using instrumental variables: Review of methods for binary instruments, treatments, and outcomes," *Journal of the American Statistical Association*, 113, 933–947.

van Kippersluis, H. and C. A. Rietveld (2017): "Pleiotropy-robust Mendelian randomization," *International Journal of Epidemiology*, 47, 1279–1288.

——— (2018): "Beyond plausibly exogenous," *The Econometrics Journal*, 21, 316–331.

A Proofs for section 3

Proof of proposition 1. Note that we assumed A1–A2 hold since they depend on observables only. Suppose equation (2) holds for all $m, \ell \in \{1, \ldots, L\}$. We will construct a joint distribution $(Y, X, Z, \widetilde{U})$ and a parameter $\widetilde{\beta}$ consistent with the data, equation (1), and assumptions A3–A4.

By the relevance assumption A1, there exists an ℓ such that $\text{cov}(X, Z_\ell) \neq 0$. Let $\widetilde{\beta} = \text{cov}(Y, Z_\ell)/\text{cov}(X, Z_\ell)$. Let $\widetilde{U} = Y - X\widetilde{\beta}$. For every $m \in \{1, \ldots, L\}$,

$$\begin{aligned}
\text{cov}(\widetilde{U}, Z_m) &= \text{cov}(Y, Z_m) - \text{cov}(X, Z_m)\widetilde{\beta} \\
&= \text{cov}(Y, Z_m) - \text{cov}(X, Z_m)\frac{\text{cov}(Y, Z_\ell)}{\text{cov}(X, Z_\ell)} \\
&= \frac{\text{cov}(Y, Z_m)\,\text{cov}(X, Z_\ell) - \text{cov}(Y, Z_\ell)\,\text{cov}(X, Z_m)}{\text{cov}(X, Z_\ell)} \\
&= 0.
\end{aligned}$$

Thus A3 holds. A4 holds by definition of $\widetilde{U}$. Thus the model is not falsified.

Next suppose the model is not falsified. Then there exists a joint distribution of (Y, X, Z, U) and a value β consistent with the model assumptions, equation (1), and the data. By A3–A4, we have

$$\text{cov}(Y, Z_\ell) = \beta\,\text{cov}(X, Z_\ell) \tag{8}$$

for all $\ell \in \{1, \ldots, L\}$. Suppose $\beta = 0$. Then $\text{cov}(Y, Z_\ell) = 0$ for all ℓ, and hence equation (2) holds for all $m, \ell \in \{1, \ldots, L\}$.

Suppose $\beta \neq 0$. Then multiplying equation (8) for ℓ by equation (8) for m gives

$$\text{cov}(Y, Z_\ell) \times (\beta\,\text{cov}(X, Z_m)) = (\beta\,\text{cov}(X, Z_\ell)) \times \text{cov}(Y, Z_m).$$

Divide by β to see that equation (2) holds for all $m, \ell \in \{1, \ldots, L\}$. $\square$

Proof of theorem 1. First we show that any value of β consistent with the model must lie in $\mathcal{B}(\delta)$. By the outcome equation (1) and the instrument exogeneity A3,

$$\text{cov}(Z, Y) = \text{cov}(Z, X)\beta + \text{var}(Z)\gamma.$$

By A2,

$$\gamma = \text{var}(Z)^{-1}(\text{cov}(Z, Y) - \text{cov}(Z, X)\beta).$$

Since $-\delta \leq \gamma \leq \delta$ (component-wise) by A4$'$, we have $\beta \in \mathcal{B}(\delta)$.

Next we show that $\mathcal{B}(\delta)$ is sharp. Let $b \in \mathcal{B}(\delta)$. Define

$$\gamma = \text{var}(Z)^{-1}(\text{cov}(Z, Y) - \text{cov}(Z, X)b).$$

Then γ satisfies A4$'$ by definition of $\mathcal{B}(\delta)$. Next, define $\widetilde{U} \equiv Y - X'b - Z'\gamma$. Then

$$\begin{aligned}\operatorname{cov}(Z, \widetilde{U}) &= \operatorname{cov}(Z, Y) - \operatorname{cov}(Z, X)b - \operatorname{var}(Z)\gamma \\ &= \operatorname{cov}(Z, Y) - \operatorname{cov}(Z, X)b - \operatorname{var}(Z)\operatorname{var}(Z)^{-1}(\operatorname{cov}(Z, Y) - \operatorname{cov}(Z, X)b) \\ &= 0.\end{aligned}$$

Thus A3 holds. Hence $\mathcal{B}(\delta)$ is sharp. That the model is falsified if and only if this set is empty follows by the definition of the (sharp) identified set. □

Proof of corollary 1. Write the identified set from theorem 1 as

$$\begin{aligned}\mathcal{B}(\delta) &= \{b \in \mathbb{R} : -\delta \leq \psi - b\pi \leq \delta\} \\ &= \{b \in \mathbb{R} : \psi_\ell - \delta_\ell \leq b\pi_\ell \leq \psi_\ell + \delta_\ell,\ \ell = 1, \ldots, L\}.\end{aligned}$$

Equation (4) follows immediately by considering the cases $\pi_\ell = 0$, $\pi_\ell < 0$, and $\pi_\ell > 0$ separately. □

Proof of lemma 1. Without loss of generality, let $\ell = 1$. The result for $\ell \neq 1$ can be obtained by permuting the components of the vector Z. Then $\widetilde{X} = (X, Z_2, \ldots, Z_L)$. Hence

$$\operatorname{cov}(Z, \widetilde{X}_1) = \begin{pmatrix} \operatorname{cov}(Z_1, X) & \operatorname{cov}(Z_1, Z_{-1}) \\ \operatorname{cov}(Z_{-1}, X) & \operatorname{var}(Z_{-1}) \end{pmatrix}.$$

By block matrix inversion, the first row of $\operatorname{cov}(Z, \widetilde{X}_1)^{-1}$ is

$$\begin{aligned}&e_1' \operatorname{cov}(Z, \widetilde{X}_1)^{-1} = \\ &\begin{pmatrix} (\operatorname{cov}(Z_1, X) - \operatorname{cov}(Z_1, Z_{-1})\operatorname{var}(Z_{-1})^{-1}\operatorname{cov}(Z_{-1}, X))^{-1} \\ -(\operatorname{cov}(Z_1, X) - \operatorname{cov}(Z_1, Z_{-1})\operatorname{var}(Z_{-1})^{-1}\operatorname{cov}(Z_{-1}, X))^{-1}\operatorname{cov}(Z_1, Z_{-1})\operatorname{var}(Z_{-1})^{-1} \end{pmatrix}'.\end{aligned}$$

Let $\widetilde{Z}_1 = Z_1 - \operatorname{cov}(Z_1, Z_{-1})\operatorname{var}(Z_{-1})^{-1}Z_{-1}$ be the population residual from a regression of Z_1 on Z_{-1}. Then

$$\begin{aligned}e_1' \operatorname{cov}(Z, \widetilde{X}_1)^{-1}\operatorname{cov}(Z, Y) &= \frac{\operatorname{cov}(Z_1, Y) - \operatorname{cov}(Z_1, Z_{-1})\operatorname{var}(Z_{-1})^{-1}\operatorname{cov}(Z_{-1}, Y)}{\operatorname{cov}(Z_1, X) - \operatorname{cov}(Z_1, Z_{-1})\operatorname{var}(Z_{-1})^{-1}\operatorname{cov}(Z_{-1}, X)} \\ &= \frac{\operatorname{cov}(\widetilde{Z}_1, Y)}{\operatorname{var}(\widetilde{Z}_1)} \Bigg/ \frac{\operatorname{cov}(\widetilde{Z}_1, X)}{\operatorname{var}(\widetilde{Z}_1)} \\ &= \frac{\psi_1}{\pi_1}.\end{aligned}$$

The last line follows by the partitioned regression formula. □

We use the following lemma in the proofs of proposition 2 and theorem 2. It says that the identified set for β is a singleton at any point δ in the set FF defined in equation (5).

Lemma 2. Suppose A1–A3 hold. Suppose $K = 1$. Let

$$b \in \left[\min_{\ell=1,\ldots,L:\pi_\ell \neq 0} \frac{\psi_\ell}{\pi_\ell}, \max_{\ell=1,\ldots,L:\pi_\ell \neq 0} \frac{\psi_\ell}{\pi_\ell} \right].$$

Define $\delta(b) = (|\psi_1 - b\pi_1|, \ldots, |\psi_L - b\pi_L|)$. Then $\mathcal{B}(\delta(b)) = \{b\}$.

Proof of lemma 2. We have

$$\begin{aligned}
\mathcal{B}(\delta(b)) &= \bigcap_{\ell=1,\ldots,L:\pi_\ell \neq 0} \left[\frac{\psi_\ell}{\pi_\ell} - \frac{|\psi_\ell - b\pi_\ell|}{|\pi_\ell|}, \; \frac{\psi_\ell}{\pi_\ell} + \frac{|\psi_\ell - b\pi_\ell|}{|\pi_\ell|} \right] \\
&= \left(\bigcap_{\ell=1,\ldots,L:\psi_\ell \geq b\pi_\ell, \pi_\ell \neq 0} \left[\frac{\psi_\ell}{\pi_\ell} - \left| \frac{\psi_\ell}{\pi_\ell} - b \right|, \frac{\psi_\ell}{\pi_\ell} + \left| \frac{\psi_\ell}{\pi_\ell} - b \right| \right] \right) \\
&\quad \bigcap \left(\bigcap_{\ell=1,\ldots,L:\psi_\ell < b\pi_\ell, \pi_\ell \neq 0} \left[\frac{\psi_\ell}{\pi_\ell} - \left| \frac{\psi_\ell}{\pi_\ell} - b \right|, \frac{\psi_\ell}{\pi_\ell} + \left| \frac{\psi_\ell}{\pi_\ell} - b \right| \right] \right) \\
&= \left(\bigcap_{\ell=1,\ldots,L:\psi_\ell \geq b\pi_\ell, \pi_\ell \neq 0} \left[b, 2\frac{\psi_\ell}{\pi_\ell} - b \right] \right) \bigcap \left(\bigcap_{\ell=1,\ldots,L:\psi_\ell < b\pi_\ell, \pi_\ell \neq 0} \left[2\frac{\psi_\ell}{\pi_\ell} - b, b \right] \right) \\
&= \{b\}.
\end{aligned}$$

The first line follows by equation (4) and the definition of $\delta(b)$. The remaining lines follow by considering two cases so that we can eliminate the absolute values. □

Proof of proposition 2. Let FF denote the true falsification frontier from definition 1. Let

$$\text{FF}^{\text{guess}} = \left\{ \delta \in \mathbb{R}^L_{\geq 0} : \delta_\ell = |\psi_\ell - b\pi_\ell|, \; \ell = 1, \ldots, L, \; b \in \left[\min_{\ell=1,\ldots,L:\pi_\ell \neq 0} \frac{\psi_\ell}{\pi_\ell}, \max_{\ell=1,\ldots,L:\pi_\ell \neq 0} \frac{\psi_\ell}{\pi_\ell} \right] \right\}.$$

We will show $\text{FF} = \text{FF}^{\text{guess}}$. We split the proof in three parts. The first two parts together show that $\text{FF}^{\text{guess}} \subseteq \text{FF}$. The third part shows that $\text{FF}^{\text{guess}} \supseteq \text{FF}$.

1. We first show that if $\delta \in \text{FF}^{\text{guess}}$, then the identified set $\mathcal{B}(\delta)$ is not empty. This follows immediately from lemma 2.

2. We next show that $\delta' < \delta$ for $\delta \in \text{FF}^{\text{guess}}$ implies that $\mathcal{B}(\delta')$ is empty. So let $\delta' < \delta$ where $\delta \in \text{FF}^{\text{guess}}$ and $\delta' \geq 0$. Consider two cases.

(a) First suppose $\delta'_\ell < \delta_\ell$ for some ℓ such that $\pi_\ell = 0$. By the definition of FF^{guess}, $\delta_\ell = |\psi_\ell|$. Note that $\delta_\ell > \delta'_\ell \geq 0$ implies $\psi_\ell \neq 0$. If $\psi_\ell > 0$ then $\psi_\ell - \delta'_\ell > 0$. So $0 \notin [\psi_\ell - \delta'_\ell, \psi_\ell + \delta'_\ell]$. Hence $B_\ell(\delta'_\ell) = \emptyset$ by equation (4). The case for $\psi_\ell < 0$ is similar. Thus in this case we must have $\mathcal{B}(\delta') = \emptyset$.

(b) Next suppose $\delta'_\ell < \delta_\ell$ for some ℓ such that $\pi_\ell \neq 0$. $\delta' < \delta$ implies that $\mathcal{B}(\delta') \subseteq \mathcal{B}(\delta)$. By lemma 2, $\mathcal{B}(\delta) = \{b^*\}$ for some value b^*. Thus it suffices to show that $b^* \notin \mathcal{B}(\delta')$. That will imply that $\mathcal{B}(\delta') = \emptyset$.

To show that $b^* \notin \mathcal{B}(\delta')$ it suffices to show that $b^* \notin B_\ell(\delta')$ for some ℓ, since $\mathcal{B}(\delta')$ is the intersection of these sets over all ℓ's, by corollary 1. From that corollary we have

$$B_\ell(\delta') = \left[\frac{\psi_\ell}{\pi_\ell} - \frac{\delta'_\ell}{|\pi_\ell|}, \frac{\psi_\ell}{\pi_\ell} + \frac{\delta'_\ell}{|\pi_\ell|}\right].$$

If $b^* \leq \psi_\ell/\pi_\ell$, then $b^* = \psi_\ell/\pi_\ell - \delta_\ell/|\pi_\ell| < \psi_\ell/\pi_\ell - \delta'_\ell/|\pi_\ell|$. Therefore, $b^* \notin B_\ell(\delta')$. The case where $b^* > \psi_\ell/\pi_\ell$ is analogous. Hence $b^* \notin \mathcal{B}(\delta')$.

Steps 1 and 2 together imply that $\text{FF}^{\text{guess}} \subseteq \text{FF}$.

3. Finally, we show that $\text{FF}^{\text{guess}} \supseteq \text{FF}$. We show the contrapositive: $\delta \notin \text{FF}^{\text{guess}}$ implies that $\delta \notin \text{FF}$. So let $\delta \notin \text{FF}^{\text{guess}}$. Denote

$$b_{\min} = \min_{\ell=1,\ldots,L:\pi_\ell\neq 0} \frac{\psi_\ell}{\pi_\ell} \qquad \text{and} \qquad b_{\max} = \max_{\ell=1,\ldots,L:\pi_\ell\neq 0} \frac{\psi_\ell}{\pi_\ell}.$$

There are two cases to consider.

(a) Suppose $\mathcal{B}(\delta) \subseteq [b_{\min}, b_{\max}]$. If $\mathcal{B}(\delta) = \emptyset$ then $\delta \notin \text{FF}$. So we can assume $\mathcal{B}(\delta) \neq \emptyset$. We'll show that we can find a $\delta' < \delta$ such that $\mathcal{B}(\delta') \neq \emptyset$, and hence $\delta \notin \text{FF}$. Let $b' \in \mathcal{B}(\delta) \subseteq [b_{\min}, b_{\max}]$. First, $\delta(b') \in \text{FF}^{\text{guess}}$ and $\delta \notin \text{FF}^{\text{guess}}$ imply that $\delta \neq \delta(b')$. Since $b' \in \mathcal{B}(\delta)$, $\delta_\ell(b') = |\psi_\ell - b'\pi_\ell| \leq \delta_\ell$ for all ℓ. Thus $\delta(b') < \delta$. Next, $\mathcal{B}(\delta(b')) = \{b'\} \neq \emptyset$ by lemma 2. So if we let $\delta' = \delta(b')$ then we have $\delta' < \delta$ and $\mathcal{B}(\delta') \neq \emptyset$. So $\delta \notin \text{FF}$ by definition of the falsification frontier.

(b) Suppose $\mathcal{B}(\delta)$ contains an element $b \notin [b_{\min}, b_{\max}]$. Suppose $b > b_{\max}$. Let $\delta' = \delta(b_{\max}) \in \text{FF}^{\text{guess}}$. By $\delta \notin \text{FF}^{\text{guess}}$, $\delta \neq \delta'$. If ℓ is such that $\pi_\ell = 0$, then $\delta'_\ell = |\psi_\ell| = |\psi_\ell - b\pi_\ell| \leq \delta_\ell$ by definition of $b \in \mathcal{B}_\ell(\delta_\ell)$. If ℓ is such that $\pi_\ell \neq 0$, then

$$\delta'_\ell = |\psi_\ell - b_{\max}\pi_\ell| = |\pi_\ell|\,(b_{\max} - \psi_\ell/\pi_\ell) < |\pi_\ell|\,(b - \psi_\ell/\pi_\ell) \leq \delta_\ell$$

by $b > b_{\max}$. So $\delta' \leq \delta$. Together with $\delta \neq \delta'$, we have $\delta' < \delta$. Also, $\mathcal{B}(\delta') = \{b_{\max}\} \neq \emptyset$ by lemma 2. So $\delta \notin \text{FF}$ by definition of the falsification frontier. A similar argument applies if instead we have $b < b_{\min}$.

□

Proof of theorem 2. We have

$$\begin{aligned}
\bigcup_{\delta \in \text{FF}} \mathcal{B}(\delta) &= \bigcup_{b \in \left[\min_{\ell=1,\ldots,L:\pi_\ell \neq 0} \frac{\psi_\ell}{\pi_\ell}, \max_{\ell=1,\ldots,L:\pi_\ell \neq 0} \frac{\psi_\ell}{\pi_\ell}\right]} \mathcal{B}(\delta(b)) \\
&= \bigcup_{b \in \left[\min_{\ell=1,\ldots,L:\pi_\ell \neq 0} \frac{\psi_\ell}{\pi_\ell}, \max_{\ell=1,\ldots,L:\pi_\ell \neq 0} \frac{\psi_\ell}{\pi_\ell}\right]} \{b\} \\
&= \left[\min_{\ell=1,\ldots,L:\pi_\ell \neq 0} \frac{\psi_\ell}{\pi_\ell}, \max_{\ell=1,\ldots,L:\pi_\ell \neq 0} \frac{\psi_\ell}{\pi_\ell}\right].
\end{aligned}$$

The first equality follows by the characterization of the falsification frontier in proposition 2. The second equality follows by lemma 2. □

Proof of proposition 3. It suffices to show that

$$\begin{pmatrix} \min_{\ell \in \widehat{\mathcal{L}}_{\text{rel}}} \widehat{b}_\ell \\ \max_{\ell \in \widehat{\mathcal{L}}_{\text{rel}}} \widehat{b}_\ell \end{pmatrix} \xrightarrow{p} \begin{pmatrix} \min_{\ell \in \mathcal{L}_{\text{rel}}} \psi_\ell/\pi_\ell \\ \max_{\ell \in \mathcal{L}_{\text{rel}}} \psi_\ell/\pi_\ell \end{pmatrix}.$$

We have

$$\begin{aligned}
\mathbb{P}\left(\min_{\ell \in \widehat{\mathcal{L}}_{\text{rel}}} \widehat{b}_\ell = \min_{\ell \in \mathcal{L}_{\text{rel}}} \widehat{b}_\ell\right) &\geq \mathbb{P}(\widehat{\mathcal{L}}_{\text{rel}} = \mathcal{L}_{\text{rel}}) \\
&= \mathbb{P}\left(\bigcap_{\ell:\pi_\ell = 0} \{F_\ell < C_n\} \cap \bigcap_{\ell:\pi_\ell \neq 0} \{F_\ell \geq C_n\}\right) \\
&= \mathbb{P}\left(\bigcap_{\ell:\pi_\ell = 0} \{C_n^{-1} F_\ell < 1\} \cap \bigcap_{\ell:\pi_\ell \neq 0} \{n^{-1} F_\ell - n^{-1} C_n \geq 0\}\right).
\end{aligned}$$

This probability converges to 1 as $n \to \infty$. To see that, use assumptions 2–4 to get $C_n^{-1} F_\ell \mathbb{1}(\pi_\ell = 0) = C_n^{-1} O_p(1) = o_p(1)$, which is < 1 with probability approaching 1. Similarly, $n^{-1} F_\ell \mathbb{1}(\pi_\ell \neq 0) - n^{-1} C_n \mathbb{1}(\pi_\ell \neq 0) = \kappa_\ell + o_p(1) - o(1)$, which is ≥ 0 with probability approaching 1. Thus $\min_{\ell \in \widehat{\mathcal{L}}_{\text{rel}}} \widehat{b}_\ell = \min_{\ell \in \mathcal{L}_{\text{rel}}} \widehat{b}_\ell + o_p(1)$. By consistency of $\widehat{b}_\ell$ for $\ell \in \mathcal{L}_{\text{rel}}$ (assumption 1) and continuity of the minimum function, $\min_{\ell \in \mathcal{L}_{\text{rel}}} \widehat{b}_\ell \xrightarrow{p} \min_{\ell \in \mathcal{L}_{\text{rel}}} \psi_\ell/\pi_\ell$. The same analysis applies to the maximum. □

Supplement to
“Salvaging Falsified Instrumental Variable Models”

Matthew A. Masten* Alexandre Poirier†

February 7, 2021

In this appendix we extend our analysis of the linear instrumental variable model to allow for multiple endogenous variables.

A The FF and FAS for Multiple Endogenous Variables

Theorem 1 characterizes the identified set for the vector of coefficients on the endogenous variables, as a function of the exclusion restriction relaxation. Our subsequent characterizations of the falsification frontier and the falsification adaptive set, however, restricted attention to the case with just one endogenous variable—see proposition 2 and theorem 2. In this section, we extend these two results to the general case with $K \geq 1$ endogenous variables. These results can also be used if, for example, a single endogenous variable has interactions with covariates or if the outcome equation is nonlinear in this variable.

In this general case we assume all instruments are relevant for simplicity. To state our new assumption, we consider submatrices of Π. Let $\mathcal{L} \subseteq \{1, \ldots, L\}$. Let $\Pi_{\mathcal{L}}$ be the $|\mathcal{L}| \times K$ submatrix of Π formed by removing all rows $\ell \notin \mathcal{L}$. Let π_ℓ' denote the ℓth row of the matrix Π. We strengthen and generalize assumption A1 as follows.

Assumption A1′ (Relevance). The following hold:

1. For all $\mathcal{L} \subseteq \{1, \ldots, L\}$ with $|\mathcal{L}| = K$, $\Pi_{\mathcal{L}}$ has full rank.

2. For all $\mathcal{L} \subseteq \{1, \ldots, L\}$ such that $|\mathcal{L}| = K + 1$, $\{\pi_\ell : \ell \in \mathcal{L}\}$ are affinely independent. That is, for all $\mathcal{L} = \{\ell_1, \ldots, \ell_{K+1}\}$,

$$\begin{pmatrix} 1 & 1 & \ldots & 1 \\ \pi_{\ell_1} & \pi_{\ell_2} & \ldots & \pi_{\ell_{K+1}} \end{pmatrix}$$

*Department of Economics, Duke University, matt.masten@duke.edu
†Department of Economics, Georgetown University, alexandre.poirier@georgetown.edu

has full rank.

A1.1′ implies that any set $\mathcal{L}$ of K instruments uniquely define the coefficients $\beta_{\mathcal{L}}^{\text{2SLS}} = \Pi_{\mathcal{L}}^{-1}\psi_{\mathcal{L}}$, where $\psi_{\mathcal{L}}$ equals the subvector of ψ after removing all components $\ell \notin \mathcal{L}$. $\beta_{\mathcal{L}}^{\text{2SLS}}$ is the population 2SLS coefficient on X using $Z_{\mathcal{L}}$ as excluded instruments and $Z_{-\mathcal{L}}$ as controls. Here we partition $Z = (Z_{\mathcal{L}}, Z_{-\mathcal{L}})$ based on the indices in $\mathcal{L}$.

A1.2′ means that there does not exist a hyperplane that passes through all of the π_ℓ vectors. It is equivalent to linear independence of $(\pi_L - \pi_1, \ldots, \pi_2 - \pi_1)$.

For $|\mathcal{L}| = K+1$, let

$$\mathcal{P}_{\mathcal{L}} = \text{conv}\big(\{\beta_{\mathcal{L}\setminus\{\ell\}}^{\text{2SLS}} : \ell \in \mathcal{L}\}\big).$$

This is the convex hull of $K+1$ just-identified 2SLS estimands in $\mathbb{R}^K$. We show that the falsification frontier and the falsification adaptive set can be constructed from $\mathcal{P}_{\mathcal{L}}$.

Proposition S1. Suppose A1′, A2, and A3 hold. Suppose the joint distribution of (Y, X, Z) is known. Then the falsification frontier is the set

$$\text{FF} = \left\{\delta \in \mathbb{R}_{\geq 0}^L : \delta_\ell = |\psi_\ell - \pi_\ell' b|,\ b \in \mathcal{P}_{\mathcal{L}},\ \mathcal{L} \subseteq \{1, \ldots, L\},\ |\mathcal{L}| = K+1\right\}.$$

Theorem S1. Suppose A1′, A2, and A3 hold. Suppose the joint distribution of (Y, X, Z) is known. Let

$$\mathcal{P} = \bigcup_{\mathcal{L} \subseteq \{1,\ldots,L\} : |\mathcal{L}| = K+1} \mathcal{P}_{\mathcal{L}}.$$

Then $\mathcal{P}$ is the falsification adaptive set.

Like the $K = 1$ case (theorem 2), $\mathcal{P}$ can be computed by running a variety of 2SLS regressions. Unlike that case, however, $\mathcal{P}$ is generally not convex, even though each $\mathcal{P}_{\mathcal{L}}$ is convex. Nonetheless, we are often only interested in linear functionals of the coefficient vector β. For example, we often care about just one component of β. The following corollary shows that the falsification adaptive set for a linear functional of β again has a simple form. For this result, let

$$\text{FAS}^* = \text{conv}\left(\{\beta_{\mathcal{L}}^{\text{2SLS}} : \mathcal{L} \subseteq \{1, \ldots, L\}, |\mathcal{L}| = K\}\right) \tag{S1}$$

denote the convex hull of the set of all just-identified 2SLS estimands.

Corollary S1. Suppose A1′, A2, and A3 hold. Suppose the joint distribution of (Y, X, Z) is known. Then FAS^* contains the falsification adaptive set for β. Moreover, for any $\alpha \in \mathbb{R}^K$

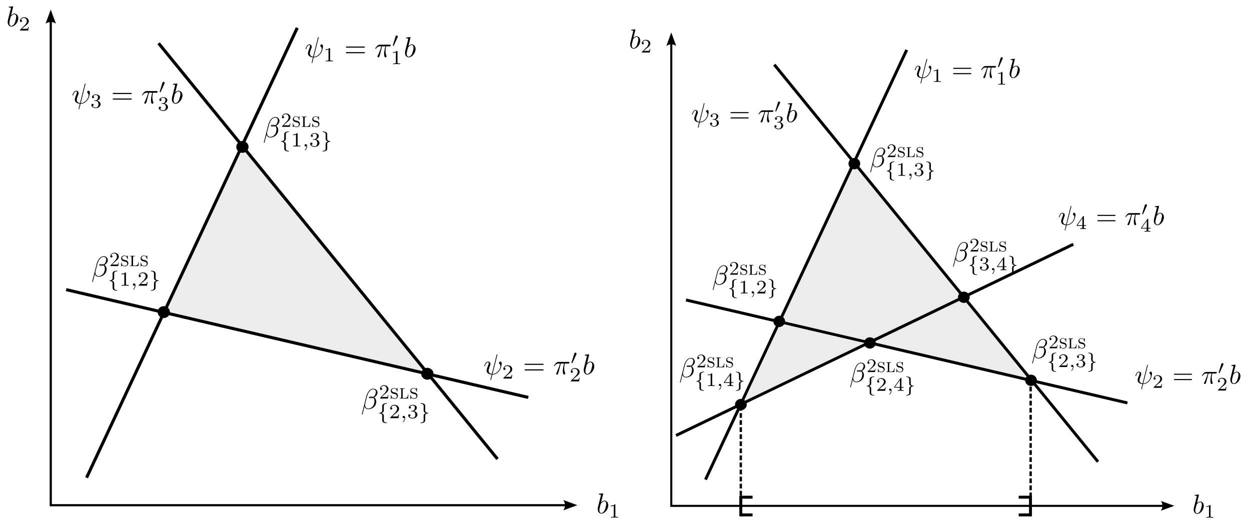

Figure 1: Example with $K = 2$ endogenous variables. Left: $L = 3$ instruments. Right: $L = 4$ instruments. In both plots, the falsification adaptive set for (β_1, β_2) is the shaded region. In the right plot, the falsification adaptive set for β_1 is shown as the projection onto the first component.

the falsification adaptive set for $\alpha'\beta$ is

$$\left[\min_{\mathcal{L} \subseteq \{1,\ldots,L\}: |\mathcal{L}|=K} \alpha' \beta^{2SLS}_{\mathcal{L}}, \max_{\mathcal{L} \subseteq \{1,\ldots,L\}: |\mathcal{L}|=K} \alpha' \beta^{2SLS}_{\mathcal{L}} \right].$$

This result shows that the FAS characterized in theorem S1 is contained in the simpler set FAS*. It also shows that we can simply cycle through all possible just identified models, compute the corresponding 2SLS estimand, take the convex hull, and project it onto one component to get the FAS for that component.

To illustrate these results, consider the two endogenous variables ($K = 2$) and three instruments ($L = 3$) case. Hence we have $L = K + 1$. Consider the left plot in figure 1. This plot shows possible values (b_1, b_2) of the coefficients on X. The exclusion restriction from instrument ℓ imposes a single linear constraint $\psi_\ell = \pi_\ell' b$. These constraints are simply lines in $\mathbb{R}^2$. Since there are three instruments, there are three constraints. When these three lines do not intersect at a common point, the baseline model is falsified. This case is shown in the figure. Suppose we drop the exclusion restriction for instrument ℓ. Then two linear constraints remain, β is point identified, and it equals the intersection point $\beta^{2SLS}_{\{1,2,3\}\setminus\{\ell\}}$. Repeating this for $\ell \in \{1, 2, 3\}$ and taking the convex hull yields the falsification adaptive set, which is shown as the shaded triangular region.

The right plot in figure 1 illustrates the $L > K + 1$ case. Here we have $K = 2$ and $L = 4$. There are 6 different just-identified 2SLS estimands. The falsification adaptive set is no longer convex. Nonetheless, the projection of the convex hull of *all* just-identified 2SLS estimands onto the first component gives the falsification adaptive set for β_1. Moreover, this projection can be computed by simply taking the largest and smallest values of β_1 among the just-identified 2SLS estimands, as shown in corollary S1.

B Proofs for section A

We next present a sequence of lemmas that lead to the proofs of the results in section A. Here we omit proofs of some of the more straightforward lemmas, but full proofs are available in appendix K of Masten and Poirier (2020).

We begin by showing a basic geometric fact about the set FAS* when $L = K + 1$. Here and elsewhere we use the notation $\beta_{-\ell}^{\text{2SLS}} = \beta_{\{1,\ldots,L\}\setminus\{\ell\}}^{\text{2SLS}}$.

Lemma S1. Suppose A1′, A2, and A3 hold. Suppose $L = K + 1$. Then exactly one of the following holds:

1. $\beta_{-\ell}^{\text{2SLS}} = \beta_{-\ell'}^{\text{2SLS}}$ for all $\ell, \ell' \in \{1, \ldots, L\}$.

2. $\pi_\ell' \beta_{-\ell}^{\text{2SLS}} \neq \psi_\ell$ for all $\ell \in \{1, \ldots, L\}$.

The next lemma shows that, when $L = K + 1$ and FAS* is not a singleton, we can write any element of $\mathbb{R}^K$ as a weighted sum of our just-identified 2SLS estimands.

Lemma S2. Suppose A1′, A2, and A3 hold. Suppose $L = K + 1$. Assume that FAS* is not a singleton. Then for any $b \in \mathbb{R}^K$ there exist weights $w_1(b), \ldots, w_L(b)$ such that $\sum_{\ell=1}^L w_\ell(b) = 1$ and

$$b = \sum_{\ell=1}^{L} w_\ell(b) \beta_{-\ell}^{\text{2SLS}}.$$

Define $\delta_\ell(b) = |\psi_\ell - \pi_\ell' b|$ for all $\ell = 1, \ldots, L$. Let $\delta(b) = (\delta_1(b), \ldots, \delta_L(b))$. We next show that, in the $L = K+1$ case, the identified set for β is a singleton for $\delta = \delta(b)$, and $b \in$ FAS*.

Lemma S3. Suppose A1′, A2, and A3 hold. Suppose $L = K + 1$. Let $b \in$ FAS*. Then $\mathcal{B}(\delta(b)) = \{b\}$.

Proof of lemma S3. By lemma S1, there are two cases to consider: FAS* is either a singleton or a nondegenerate simplex in $\mathbb{R}^K$.

Case 1. Suppose FAS* $= \{b\}$ is a singleton. By the definition of FAS*, this implies that $b = \beta_{\mathcal{L}}^{\text{2SLS}}$ for any $\mathcal{L} \subseteq \{1, \ldots, L\}$ with $|\mathcal{L}| = K$. It also implies $\delta(b) = 0_L$. Therefore $\mathcal{B}(0_L) = \bigcap_{\ell=1}^L B_\ell(0) = \bigcap_{\mathcal{L}:|\mathcal{L}|=K} \left(\bigcap_{\ell\in\mathcal{L}} B_\ell(0)\right) = \{b\}$ by $\bigcap_{\ell\in\mathcal{L}} B_\ell(0) = \{b\}$.

Case 2. Suppose FAS* is not a singleton. Then $\pi_\ell' \beta_{-\ell}^{\text{2SLS}} \neq \psi_\ell$ for all $\ell \in \{1, \ldots, L\}$. We prove equality of sets by showing that both directions of set inclusion hold.

Step 1 ($\supseteq$). First we show that $\mathcal{B}(\delta(b)) \supseteq \{b\}$. By definition of $\delta_\ell(\cdot)$, $\psi_\ell - \pi_\ell' b \in [-\delta_\ell(b), \delta_\ell(b)]$ for all ℓ. Thus, by the characterization of $\mathcal{B}(\cdot)$ in theorem 1, $b \in \mathcal{B}(\delta(b))$.

Step 2 ($\subseteq$). Next we show that $\mathcal{B}(\delta(b)) \subseteq \{b\}$. First suppose $\delta(b) = 0_L$. In this case the baseline model is not falsified and FAS* is a singleton. This is a contradiction. So we must have $\delta_\ell(b) > 0$ for some ℓ.

We will show that any element $b^* \neq b$ is not in $\mathcal{B}(\delta(b))$. The set FAS* is a polytope. Consider its alternative half-space representation. The half-spaces correspond to one side of the hyperplanes $B_\ell(0)$. Formally, write

$$\text{FAS}^* = \bigcap_{\ell=1}^{L} \{\tilde{b} \in \mathbb{R}^K : \psi_\ell - \pi_\ell' \tilde{b} \leq 0\}. \tag{S2}$$

We assume without loss of generality that all L inequalities go in the same direction. This is because $\psi_\ell - \pi_\ell' \tilde{b} \geq 0$ can be rewritten as $-\psi_\ell - (-\pi_\ell' \tilde{b}) \leq 0$, which is equivalent to replacing instrument Z_ℓ with $-Z_\ell$. Neither the estimands $\beta_{\mathcal{L}}^{\text{2SLS}}$ nor the set FAS* are affected by these sign normalizations.

Noting that $\mathcal{B}(\delta)$ is an intersection of half spaces and evaluating it at $\delta(b)$ gives

$$\begin{aligned}
\mathcal{B}(\delta(b)) &= \bigcap_{\ell=1}^{L} \left\{ \tilde{b} \in \mathbb{R}^K : -\delta_\ell(b) \leq \psi_\ell - \pi_\ell' \tilde{b} \leq \delta_\ell(b) \right\} \\
&= \left(\bigcap_{\ell=1}^{L} \{\tilde{b} \in \mathbb{R}^K : \psi_\ell - \pi_\ell' \tilde{b} \geq -|\psi_\ell - \pi_\ell' b|\} \right) \bigcap \left(\bigcap_{\ell=1}^{L} \{\tilde{b} \in \mathbb{R}^K : \psi_\ell - \pi_\ell' \tilde{b} \leq |\psi_\ell - \pi_\ell' b|\} \right) \\
&\equiv \mathcal{P}_1(b) \cap \mathcal{P}_2(b).
\end{aligned}$$

To complete this proof, it suffices to show $b^* \notin \mathcal{P}_1(b)$. By lemma S2, any element in $\mathbb{R}^K$ can be written as a linear combination of the L different just-identified 2SLS estimands. In particular, we can write b^* in this way:

$$b^* = \sum_{\ell=1}^{L} w_\ell(b^*) \beta_{-\ell}^{\text{2SLS}}$$

where $w_\ell(b^*)$ are weights that sum to one, $\sum_{\ell=1}^{L} w_\ell(b^*) = 1$.

Since $b \in \text{FAS}^*$, $\psi_\ell - \pi_\ell' b \leq 0$ for all ℓ. This follows directly from our half-space representation of FAS^*. Thus $-|\psi_\ell - \pi_\ell' b| = \psi_\ell - \pi_\ell' b$ for all ℓ. Hence

$$\mathcal{P}_1(b) = \bigcap_{\ell=1}^{L} \{\tilde{b} \in \mathbb{R}^K : \psi_\ell - \pi_\ell' \tilde{b} \geq -|\psi_\ell - \pi_\ell' b|\} = \bigcap_{\ell=1}^{L} \{\tilde{b} \in \mathbb{R}^K : \pi_\ell'(\tilde{b} - b) \leq 0\}.$$

So $b^* \in \mathcal{P}_1(b)$ if and only if $\pi_\ell'(b^* - b) \leq 0$ for all ℓ. Focus on just one ℓ for a moment. Then

$$\begin{aligned}
\pi_\ell'(b^* - b) &= \sum_{s=1}^{L} (w_s(b^*) - w_s(b)) \pi_\ell' \beta_{-s}^{\text{2SLS}} \\
&= \sum_{s \neq \ell} (w_s(b^*) - w_s(b)) \psi_\ell + (w_\ell(b^*) - w_\ell(b)) \pi_\ell' \beta_{-\ell}^{\text{2SLS}} \\
&= (\psi_\ell - \pi_\ell' \beta_{-\ell}^{\text{2SLS}}) \sum_{s \neq \ell} (w_s(b^*) - w_s(b)).
\end{aligned}$$

The first line follows from lemma S2. The second follows from $\psi_\ell = \pi_\ell' \beta_{-s}^{\text{2SLS}}$ when $s \neq \ell$ by the definition of these 2SLS estimands. The third follows from the difference in weights summing to zero.

Next notice that $\psi_\ell - \pi_\ell' \beta_{-\ell}^{\text{2SLS}} < 0$. This follows from $\beta_{-\ell}^{\text{2SLS}} \in \text{FAS}^*$, the half-space representation of FAS^*, and since FAS^* is a nondegenerate simplex. Suppose by way of contradiction that $b^* \in \mathcal{P}_1(b)$. Then $\pi_\ell'(b^* - b) \leq 0$ for all ℓ. We've just seen that this implies $\sum_{s \neq \ell} (w_s(b^*) - w_s(b)) \geq 0$ for all ℓ. But now we have

$$0 = \sum_{s=1}^{L} (w_s(b^*) - w_s(b)) = \sum_{s \neq \ell} (w_s(b^*) - w_s(b)) + (w_\ell(b^*) - w_\ell(b)).$$

Thus $w_\ell(b^*) - w_\ell(b) = \sum_{s \neq \ell} (w_s(b) - w_s(b^*)) \leq 0$ for all ℓ. Since $w_\ell(b^*) - w_\ell(b)$ sums to zero, $w_\ell(b^*) = w_\ell(b)$ for all ℓ. This implies $b^* = b$, a contradiction. Thus $b^* \notin \mathcal{P}_1(b)$. □

Lemma S4. Suppose A1′, A2, and A3 hold. Suppose $L \geq K + 1$. Let $b \in \mathcal{P}$. Then $\mathcal{B}(\delta(b)) = \{b\}$.

Proof of lemma S4. We prove set equality by showing that both directions of set inclusion hold as in the proof of lemma S1.

Step 1 ($\supseteq$). The proof of this step from lemma S3 applies without modification.

Step 2 ($\subseteq$). Since $b \in \mathcal{P}$ there is some $\mathcal{L} \subseteq \{1, \ldots, L\}$ with $|\mathcal{L}| = K + 1$ such that $b \in \mathcal{P}_{\mathcal{L}}$. Let

$$\mathcal{B}_{\mathcal{L}}(\delta) = \bigcap_{\ell \in \mathcal{L}} B_\ell(\delta).$$

By lemma S3, $\mathcal{B}_{\mathcal{L}}(\delta(b)) = \{b\}$. By definition, $\mathcal{B}_{\mathcal{L}}(\delta) \supseteq \mathcal{B}(\delta)$. Thus $\mathcal{B}(\delta(b)) \subseteq \{b\}$. □

The following variation on Farkas' lemma (for example, corollary 22.3.1 on page 200 of Rockafellar 1970; Border 2019 provides an extensive discussion) is helpful.

Lemma S5 (Variation on Farkas' Lemma)**.** Let $x_1, \ldots, x_n \in \mathbb{R}^K$. Then $0_K \notin \text{conv}(\{x_1, \ldots, x_n\})$ if and only if there exists a $p \in \mathbb{R}^K$ such that $p'x_i > 0$ for all $i = 1, \ldots, n$.

The next two technical lemmas are used in the proof of lemma S8, which is then used in the proof of lemma S9.

Lemma S6. Suppose A1$'$ holds. Suppose $L = K + 1$. Suppose FAS* is not a singleton. Then $\psi_\ell - \pi_\ell' b = w_\ell(b)(\psi_\ell - \pi_\ell' \beta_{-\ell}^{\text{2SLS}})$ for all $\ell = 1, \ldots, L$.

Lemma S7. Suppose A1$'$ holds. Suppose $L = K + 1$. Suppose FAS* is not a singleton. Without loss of generality (see equation (S2) and the surrounding discussion), write

$$\text{FAS}^* = \bigcap_{\ell=1}^{L} \{b \in \mathbb{R}^K : \psi_\ell - \pi_\ell' b \leq 0\}.$$

Then there are no $b \in \mathbb{R}^K$ such that $\psi_\ell - \pi_\ell' b \geq 0$ for all $\ell = 1, \ldots, L$.

Lemma S8. Suppose A1$'$, A2, and A3 hold. Suppose $L = K + 1$. Then there exists a normalization of the hyperplanes $\{b \in \mathbb{R}^K : \pi_\ell' b = \psi_\ell\}$ with $\psi_\ell \geq 0$ and such that

$$0_K \in \text{conv}(\{\pi_\ell : \ell = 1, \ldots, L\}) \qquad \Rightarrow \qquad 0_K \in \text{conv}(\{\beta_{-\ell}^{\text{2SLS}} : \ell = 1, \ldots, L\}).$$

Lemma S9. Suppose A1$'$, A2, and A3 hold. Suppose $L \geq K + 1$. Let $b \notin \mathcal{P}$. Then there exists a $\delta' < \delta(b)$ such that $\mathcal{B}(\delta') \neq \emptyset$.

Proof of lemma S9. Without loss of generality, suppose $b = 0_K$. This follows since we can simply translate our coordinate system so that the origin is at b. Put differently, we map all $x \in \mathbb{R}^K$ to $x - b$. Throughout this proof, we also use a normalization from lemma S8 where $\psi_\ell \geq 0$ for all ℓ. Next, there are two cases to consider.

Case 1. Suppose $\delta_\ell(b) = |\psi_\ell - \pi_\ell' b| = \psi_\ell > 0$ for all ℓ. Since $b \notin \mathcal{P}$, $b \notin \mathcal{P}_{\mathcal{L}} = \text{conv}(\{\beta_{\mathcal{L}\setminus\{l\}}^{\text{2SLS}} : \ell \in \mathcal{L}\})$ for any $\mathcal{L}$ with $|\mathcal{L}| = K + 1$. Since $b = 0_K$, lemma S8 implies that $0_K \notin \text{conv}(\{\pi_\ell : \ell \in \mathcal{L}\})$. This holds for any set $\mathcal{L}$ such that $|\mathcal{L}| = K + 1$. This implies that $0_K \notin \text{conv}(\{\pi_\ell : \ell = 1, \ldots, L\})$. To see this, assume $0_K \in \text{conv}(\{\pi_\ell : \ell = 1, \ldots, L\})$. By Caratheodory's theorem, (e.g., chapter 17 of Rockafellar 1970) 0_K is then in the convex hull of a $(K+1)$-element subset of $\{\pi_\ell : \ell = 1, \ldots, L\}$, That is, $0_K \in \text{conv}(\{\pi_\ell : \ell \in \mathcal{L}\})$ for some $\mathcal{L}$ with $|\mathcal{L}| = K + 1$. This is a contradiction.

Since $0_K \notin \text{conv}(\{\pi_\ell : \ell = 1, \ldots, L\})$, lemma S5 implies that there exists a vector $\bar{b}$ such that $\pi_\ell' \bar{b} > 0$ for all $\ell = 1, \ldots, L$. Define $b(\varepsilon) = b + \varepsilon \bar{b}$. Since $\psi_\ell > 0$ and $\pi_\ell' \bar{b} > 0$ for all ℓ, there exists an $\bar{\varepsilon} > 0$ such that $\psi_\ell - \bar{\varepsilon} \pi_\ell' \bar{b} > 0$ for all ℓ. This implies that

$$0 < |\psi_\ell - \pi_\ell' b(\bar{\varepsilon})| = |\psi_\ell - \pi_\ell'(\bar{\varepsilon}\bar{b})| = \delta_\ell(b) - \bar{\varepsilon}\pi_\ell' \bar{b} < \delta_\ell(b)$$

by $\delta_\ell(b) = \psi_\ell$ and by $0 < \bar{\varepsilon}\pi_\ell' \bar{b} < \psi_\ell$ for all ℓ.

Let $\delta_\ell' = |\psi_\ell - \pi_\ell' b(\bar{\varepsilon})|$. We have shown that $\delta' < \delta(b)$. Finally, by our characterization of $\mathcal{B}(\cdot)$ and the definition of δ', we have $b(\bar{\varepsilon}) \in \mathcal{B}(\delta')$. Hence $\mathcal{B}(\delta') \neq \emptyset$.

Case 2. Suppose $\delta_\ell(b) = 0$ for some ℓ's. There can be at most $K - 1$ such indices, since otherwise we would have $b \in \mathcal{P}$. Let $\mathcal{L}_0$ denote the set of these indices. Since $b = 0_K$ and $\delta_\ell(b) = 0$, $\psi_\ell = 0$ for $\ell \in \mathcal{L}_0$. In this case, consider the subspace

$$\{\tilde{b} \in \mathbb{R}^K : 0 = \pi_\ell' \tilde{b}, \ \ell \in \mathcal{L}_0\}.$$

This is a linear subspace of dimension at least 1 (by $|\mathcal{L}_0| > 0$) and at most $K - 1$ (as noted earlier). Within this subspace, we can look at the remaining indices $\{1, \ldots, L\} \setminus \mathcal{L}_0$. We have $\delta_\ell(b) > 0$ for all of these indices. By restricting attention to this subspace we can thus immediately apply the analysis of case 1. □

For the next two lemmas, let

$$\text{FF}^{\text{guess}} = \{\delta \in \mathbb{R}^L_{\geq 0} : \delta_\ell = |\psi_\ell - \pi_\ell' b|, \ell = 1, \ldots, L, b \in \mathcal{P}\}$$

and let FF denote the true falsification frontier from definition 1.

Lemma S10. Suppose A1$'$, A2, and A3 hold. Suppose $L \geq K + 1$. Then $\text{FF}^{\text{guess}} \subseteq \text{FF}$.

Proof of lemma S10. Recall the definition $\delta_\ell(b) = |\psi_\ell - \pi_\ell' b|$. Let $\delta \in \text{FF}^{\text{guess}}$. Then, by definition, there is a $b \in \mathcal{P}$ such that $\delta_\ell(b) = \delta_\ell$ for all ℓ. Thus $\mathcal{B}(\delta) = \{b\}$ by lemma S4.

Let $\delta' < \delta(b)$. Then there is some index ℓ such that $0 < \delta_\ell' < \delta_\ell(b)$. So $\psi_\ell - \pi_\ell' b \notin [-\delta_\ell', \delta_\ell']$ and hence $b \notin B_\ell(\delta')$. This implies that $b \notin \mathcal{B}(\delta')$. But since $\mathcal{B}(\delta') \subseteq \mathcal{B}(\delta) = \{b\}$, we must have $\mathcal{B}(\delta') = \emptyset$. Hence, by the definition of the falsification frontier, $\text{FF}^{\text{guess}} \subseteq \text{FF}$. □

Lemma S11. Suppose A1$'$, A2, and A3 hold. Suppose $L \geq K + 1$. Then $\text{FF}^{\text{guess}} \supseteq \text{FF}$.

Proof of lemma S11. We will show the contrapositive: $\delta \notin \text{FF}^{\text{guess}}$ implies $\delta \notin \text{FF}$. Let $\delta \notin \text{FF}^{\text{guess}}$. There are two cases to consider.

Case 1. Suppose δ is such that $\mathcal{B}(\delta)$ contains an element $b \notin \mathcal{P}$. By lemma S9 there exists $\delta' < \delta(b)$ with $\mathcal{B}(\delta') \neq \emptyset$. Moreover, $\delta(b) \leq \delta$ by the characterization of $\mathcal{B}(\delta)$ in theorem 1. Thus $\delta \notin \text{FF}$ by the definition of the falsification frontier.

Case 2. Suppose δ is such that $\mathcal{B}(\delta) \subseteq \mathcal{P}$. If $\mathcal{B}(\delta) = \emptyset$, then $\delta \notin \text{FF}$ by definition. Therefore we let $\mathcal{B}(\delta) \neq \emptyset$. Let b' be any element of $\mathcal{B}(\delta)$. Let $\delta' = \delta(b')$. By $b' \in \mathcal{P}$, $\delta' \in \text{FF}^{\text{guess}}$ and by $\delta \notin \text{FF}^{\text{guess}}$, $\delta' \neq \delta$. Also, by $b' \in \mathcal{B}(\delta)$ we have $\delta'_\ell = |\psi_\ell - \pi'_\ell b'| \leq \delta_\ell$ for all ℓ. Together these imply $\delta' < \delta$. Moreover, we have $\mathcal{B}(\delta') = \mathcal{B}(\delta(b')) = \{b'\} \neq \emptyset$ by lemma S4. Thus $\delta \notin \text{FF}$, by definition of the falsification frontier.

All values of δ must fall in one of these two cases. Therefore $\text{FF}^{\text{guess}} \supseteq \text{FF}$. □

Proof of proposition S1. This follows directly from lemmas S10 and S11. □

Proof of theorem S1. We have

$$\bigcup_{\delta \in \text{FF}} \mathcal{B}(\delta) = \bigcup_{b \in \mathcal{P}} \mathcal{B}(\delta(b)) = \bigcup_{b \in \mathcal{P}} \{b\} = \mathcal{P}$$

by proposition S1 and lemma S4. □

To prove corollary S1, we use the following definition: Let P be a $K \times K$ idempotent matrix. Define the linear operator $p : \mathbb{R}^K \to \mathbb{R}^K$ by $p(a) = Pa$. p is called a *projection*. For $A \subseteq \mathbb{R}^K$, define the projection of the set A as $\text{proj}(A) = \{p(a) \in \mathbb{R}^K : a \in A\}$.

We use the following lemma in the proof of the corollary.

Lemma S12. $\text{proj}(\text{conv}(A)) = \text{conv}(\text{proj}(A))$.

Proof of corollary S1. The first part of this result states that $\mathcal{P} \subseteq \text{FAS}^*$. To see this, let $b \in \mathcal{P}$. Then $b \in \mathcal{P}_\mathcal{L}$ for some set of indices $\mathcal{L}$ with $|\mathcal{L}| = K + 1$. But if b is a convex combination of $\{\beta^{\text{2SLS}}_{\mathcal{L} \setminus \{\ell\}} : \ell \in \mathcal{L}\}$, then it is also a convex combination of the larger set of elements $\{\beta^{\text{2SLS}}_\mathcal{S} : \mathcal{S} \subseteq \{1, \ldots, L\}, |\mathcal{S}| = K\}$. Hence $b \in \text{FAS}^*$.

To prove the second part, consider projections defined by the $K \times K$ matrix

$$P = \begin{pmatrix} \alpha & 0_K & \cdots & 0_K \end{pmatrix}'$$

where $\alpha \in \mathbb{R}^K$. For any set $A \subseteq \mathbb{R}^K$, let $[A]_1 = \{a_1 \in \mathbb{R} : a = (a_1, \ldots, a_K) \in \mathbb{R}^K\}$. Since P maps the $2, \ldots, K$ components of any vector $a \in \mathbb{R}^K$ to zero, it suffices to show that $[\text{proj}(\mathcal{P})]_1 = [\text{proj}(\text{FAS}^*)]_1$ where

$$[\text{proj}(\text{FAS}^*)]_1 = \left[\min_{\mathcal{L} \subseteq \{1,\ldots,L\} : |\mathcal{L}| = K} \alpha' \beta^{\text{2SLS}}_\mathcal{L}, \max_{\mathcal{L} \subseteq \{1,\ldots,L\} : |\mathcal{L}| = K} \alpha' \beta^{\text{2SLS}}_\mathcal{L} \right]. \quad \text{(S3)}$$

We have

$$\text{proj}(\text{FAS}^*) = \text{proj}(\text{conv}(\mathcal{P})) = \text{conv}(\{P\beta^{\text{2SLS}}_\mathcal{L} : \mathcal{L} \subseteq \{1, \ldots, L\}, |\mathcal{L}| = K\})$$

by lemma S12. Using the specific form of the projection P now gives equation (S3).

Similarly,

$$\begin{aligned}\operatorname{proj}(\mathcal{P}) &= \bigcup_{\mathcal{L}\subseteq\{1,\ldots,L\}:|\mathcal{L}|=K+1} \operatorname{proj}\left(\operatorname{conv}\left(\left\{\beta^{\text{2SLS}}_{\mathcal{L}\setminus\{\ell\}} : \ell\in\mathcal{L}\right\}\right)\right) \\ &= \bigcup_{\mathcal{L}\subseteq\{1,\ldots,L\}:|\mathcal{L}|=K+1} \operatorname{conv}\left(\operatorname{proj}\left(\left\{\beta^{\text{2SLS}}_{\mathcal{L}\setminus\{\ell\}} : \ell\in\mathcal{L}\right\}\right)\right).\end{aligned}$$

The first line follows by commutativity of projections and unions. The second from lemma S12. Hence

$$[\operatorname{proj}(\mathcal{P})]_1 = \bigcup_{\mathcal{L}\subseteq\{1,\ldots,L\}:|\mathcal{L}|=K+1} \left[\min_{\ell\in\mathcal{L}} \alpha'\beta^{\text{2SLS}}_{\mathcal{L}\setminus\{\ell\}},\ \max_{\ell\in\mathcal{L}} \alpha'\beta^{\text{2SLS}}_{\mathcal{L}\setminus\{\ell\}}\right].$$

Thus we see that the first component of $\operatorname{proj}(\mathcal{P})$ is a union of closed intervals. If $\mathcal{L}$ and $\mathcal{L}'$ differ by at most one element, then the intervals

$$\left[\min_{\ell\in\mathcal{L}} \alpha'\beta^{\text{2SLS}}_{\mathcal{L}\setminus\{\ell\}},\ \max_{\ell\in\mathcal{L}} \alpha'\beta^{\text{2SLS}}_{\mathcal{L}\setminus\{\ell\}}\right] \quad \text{and} \quad \left[\min_{\ell\in\mathcal{L}'} \alpha'\beta^{\text{2SLS}}_{\mathcal{L}'\setminus\{\ell\}},\ \max_{\ell\in\mathcal{L}'} \alpha'\beta^{\text{2SLS}}_{\mathcal{L}'\setminus\{\ell\}}\right]$$

have non-empty intersection. Their union is therefore a closed interval. Because we take the union over all $\mathcal{L}\subseteq\{1,\ldots,L\} : |\mathcal{L}| = K+1$, we can find a sequence $(\mathcal{L}_1,\ldots,\mathcal{L}_N)$ such that

$$\left[\min_{\ell\in\mathcal{L}_n} \alpha'\beta^{\text{2SLS}}_{\mathcal{L}_n\setminus\{\ell\}},\ \max_{\ell\in\mathcal{L}_n} \alpha'\beta^{\text{2SLS}}_{\mathcal{L}_n\setminus\{\ell\}}\right] \quad \text{and} \quad \left[\min_{\ell\in\mathcal{L}_{n+1}} \alpha'\beta^{\text{2SLS}}_{\mathcal{L}_{n+1}\setminus\{\ell\}},\ \max_{\ell\in\mathcal{L}_{n+1}} \alpha'\beta^{\text{2SLS}}_{\mathcal{L}_{n+1}\setminus\{\ell\}}\right]$$

overlap for $n = 1,\ldots,N-1$ and such that $\bigcup_{n=1}^N \mathcal{L}_n = \{1,\ldots,L\}$. Thus

$$\bigcup_{\mathcal{L}\subseteq\{1,\ldots,L\}:|\mathcal{L}|=K+1} \left[\min_{\ell\in\mathcal{L}} \alpha'\beta^{\text{2SLS}}_{\mathcal{L}\setminus\{\ell\}}, \max_{\ell\in\mathcal{L}} \alpha'\beta^{\text{2SLS}}_{\mathcal{L}\setminus\{\ell\}}\right] = \left[\min_{\mathcal{L}\subseteq\{1,\ldots,L\}:|\mathcal{L}|=K} \alpha'\beta^{\text{2SLS}}_{\mathcal{L}},\ \max_{\mathcal{L}\subseteq\{1,\ldots,L\}:|\mathcal{L}|=K} \alpha'\beta^{\text{2SLS}}_{\mathcal{L}}\right].$$

Putting everything together yields $[\operatorname{proj}(\mathcal{P})]_1 = [\operatorname{proj}(\text{FAS}^*)]_1$ as desired. □

References

Border, K. C. (2019): "Alternative linear inequalities," V. 2019.10.23::09.57.

Masten, M. A. and A. Poirier (2020): "Salvaging falsified instrumental variable models," *arXiv preprint arXiv:1812.11598v3*.

Rockafellar, R. T. (1970): *Convex Analysis*, Princeton University Press.