The Falsification Adaptive Set in Linear Models with Instrumental Variables that Violate the Exogeneity or Exclusion Restriction

Nicolas Apfel[a] and Frank Windmeijer[b]

[a]School of Economics, University of Surrey, UK

[b]Dept of Statistics and Nuffield College, University of Oxford, UK

December 19, 2022

Abstract

For the classical linear model with an endogenous variable estimated by the method of instrumental variables (IVs) with multiple instruments, Masten and Poirier (2021) introduced the falsification adaptive set (*FAS*). When a model is falsified, the *FAS* reflects the model uncertainty that arises from falsification of the baseline model. It is the set of just-identified IV estimands, where each relevant instrument is considered as the just-identifying instrument in turn, whilst all other instruments are included as controls. It therefore applies to the case where the exogeneity assumption holds and invalid instruments violate the exclusion assumption only. We propose a generalized *FAS* that reflects the model uncertainty when some instruments violate the exogeneity assumption and/or some instruments violate the exclusion assumption. This *FAS* is the set of all possible just-identified IV estimands where the just-identifying instrument is relevant. There are a maximum of $k_z 2^{k_z-1}$ such estimands, where k_z is the number of instruments. If there is at least one relevant instrument that is valid in the sense that it satisfies the exogeneity and exclusion assumptions, then this generalized *FAS* is guaranteed to contain β and therefore to be the identified set for β.

Keywords: Instrumental variables, invalid instruments, falsification adaptive set
JEL Codes: C26, C52

1 Introduction

Masten and Poirier (2021) (henceforth MP) introduced the falsification adaptive set (FAS) for the classical linear model with an endogenous variable estimated by the method of instrumental variables with multiple instruments. For the model specification $Y = X\beta + \boldsymbol{Z}^T\boldsymbol{\gamma}+U$, where $\boldsymbol{Z}$ is a k_z-vector of putative instruments, the exclusion assumption implies that $\boldsymbol{\gamma} = \mathbf{0}$, and the exogeneity assumption that $\text{cov}\,(\boldsymbol{Z}, U) = \mathbf{0}$. The FAS as proposed by MP applies to the case where the baseline model is falsified due to instruments violating the exclusion assumption whilst maintaining the exogeneity assumption. We therefore denote this the FAS_{excl}. It is the set of just-identified, and hence non-falsifiable, IV estimands, where each relevant instrument is considered as the just-identifying instrument in turn, whilst all other instruments are included as controls. The FAS_{excl} is thus an expanded set compared to the baseline point estimand to account for the uncertainty due to a violation of the exclusion assumption. MP specify the falsification frontier as the set of smallest relaxations of the exclusion restriction that are not falsified. They recommend to report estimates of FAS_{excl} under the assumption that the true model lies on this frontier, as then the FAS_{excl} contains β and is therefore the identified set for β.

An instrument is valid if it satisfies both the exogeneity and exclusion assumption. As we show in Section 3, the FAS_{excl} is guaranteed to contain β if there is at least one valid and relevant instrument if the exogeneity assumption holds and invalid instruments violate the exclusion assumption only. MP argue that, mathematically, the same technical analysis can be used to relax both the exclusion assumption and exogeneity assumption. However, we show in Section 4 that the FAS_{excl} is no longer guaranteed to contain β if there is at least one valid and relevant instrument when instead the exclusion assumption holds and invalid instruments violate the exogeneity assumption only. We propose a different *FAS,* termed FAS_{exo}, for this setting. FAS_{exo} is the set of just-identified IV estimands where each relevant instrument is considered as the just-identifying instrument in turn, with all other instruments removed from the instrument set. We show that FAS_{exo} is guaranteed to contain β when at least one of the instruments is relevant and valid and invalid instruments can violate the exogeneity assumption only.

From these results, and as the main contribution of this paper, we argue in Section 5 that a FAS that properly takes into account possible violations of both the exogeneity and exclusion assumptions when the baseline model is falsified is the set of all possible just-identified IV estimands where the just-identifying instrument is relevant. This is due to the fact that when an instrument violates the exogeneity assumption it should be removed from the instrument set, and when an instrument violates the exclusion assumption it should be included as a control. Whereas FAS_{excl} and FAS_{exo} consider a maximum of k_z just-identified estimands, our generalized FAS considers a maximum of $k_z 2^{k_z-1}$ just-identified estimands. Under the assumption that an invalid instrument can either violate the exogeneity assumption or the exclusion assumption, this generalized FAS is guaranteed to contain β when there is at least one valid and relevant instrument.

In Section 5.1 we first introduce the generalized FAS for the case where there are two instruments available, as for this case it is simply the union of the FAS_{excl} and FAS_{exo}. Section 5.2 provides the details for the generalized FAS for a general number of instruments. In Section 6.1 we compare estimates of the three falsification adaptive sets for an empirical example with two instruments on the origin of gender roles taken from Alesina et al. (2013), as considered in Masten and Poirier (2020) (henceforth MP20). Section 6.2 presents estimation results for an empirical analysis of roads and trade with three instruments taken from Duranton et al. (2014), which was the main application in MP.

2 Model and Assumptions

MP consider the classical linear model. In the main text they derive the falsification adaptive set for a model with one endogenous explanatory variable and we focus on that model here. $Y(x, \boldsymbol{z})$ denotes the potential outcomes defined for values $(x, \boldsymbol{z}) \in \mathbb{R}^{k_z+1}$. It is assumed that

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

where β is an unknown constant, $\boldsymbol{\gamma}$ is an unknown k_z-vector and U is an unobserved random variable. Let X be an observed endogenous variable and let $\boldsymbol{Z}$ be an observed k_z-vector of potentially invalid instruments. Individual elements of $\boldsymbol{Z}$ are denoted Z_ℓ, $\ell = 1, \ldots, k_z$. The observed outcome is $Y = Y(X, \boldsymbol{Z})$,

$$Y = X\beta + \boldsymbol{Z}^T \boldsymbol{\gamma} + U, \tag{2}$$

and the joint distribution of $(Y, X, \boldsymbol{Z})$ is assumed known.

The exclusion assumption as detailed below is an assumption on the values of γ_ℓ. In order to state the exogeneity assumption in a similar way, we specify

$$\text{cov}\left(\boldsymbol{Z}, U\right) = \boldsymbol{\alpha}, \tag{3}$$

where $\boldsymbol{\alpha}$ is an unknown k_z-vector.

The following relevance and sufficient variation assumptions are maintained:

Assumption 1. *Relevance: The k_z-vector $cov(\boldsymbol{Z}, X) \neq \mathbf{0}$.*

Assumption 2. *Sufficient variation: The $k_z \times k_z$ matrix $\boldsymbol{\Sigma}_z := var(\boldsymbol{Z})$ is invertible.*

For all instruments to be valid, the exogeneity and exclusion assumptions need to be satisfied:

Assumption 3. *Exogeneity: $\alpha_\ell = 0$ for all $\ell \in \{1, \ldots, k_z\}$.*

Assumption 4. *Exclusion: $\gamma_\ell = 0$ for all $\ell \in \{1, \ldots, k_z\}$.*

Under model (1) and Assumptions 1-4 it follows that $\text{cov}\left(\boldsymbol{Z}, Y\right) = \text{cov}\left(\boldsymbol{Z}, X\right)\beta$ and

$$\left(\text{cov}\left(\boldsymbol{Z}, X\right)\right)^T \boldsymbol{\Sigma}_z^{-1}\left(\text{cov}\left(\boldsymbol{Z}, X\right)\right)^{-1}\left(\text{cov}\left(\boldsymbol{Z}, X\right)\right)^T \boldsymbol{\Sigma}_z^{-1}\left(\text{cov}\left(\boldsymbol{Z}, Y\right)\right) = \beta,$$

and so the two-stage least squares (2sls) estimand is equal to β. For each individual instrument, we have that $\text{cov}\left(Z_\ell, Y\right) = \text{cov}\left(Z_\ell, X\right)\beta$. Maintaining Assumptions 1 and 2, Proposition 1 in MP states that model (2) with Assumptions 3 and 4 is not falsified if and only if

$$\text{cov}\left(Z_m, Y\right)\text{cov}\left(Z_\ell, X\right) \neq \text{cov}\left(Z_\ell, Y\right)\text{cov}\left(Z_m, X\right)$$

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

As discussed in MP, if the distribution of $(Y, X, \boldsymbol{Z})$ is such that the model is falsified then this could be due to misspecification of model (2), which assumes homogeneous linear treatment effects, and/or instrument invalidity. As in MP, we maintain here model (2) and Assumptions 1 and 2 and focus on failure of instrument exogeneity or instrument exclusion as reasons for falsifying the baseline model.

We define an instrument to be valid as follows.

Definition 1. Valid instrument: An instrument Z_ℓ is a valid instrument if both the exogeneity and exclusion assumptions hold, $\alpha_\ell = \gamma_\ell = 0$.

We further make the assumption that an invalid instrument can either violate the exclusion assumption or the exogeneity assumption, but not both:

Assumption 5. *Invalid instrument: An invalid instrument violates either the exogeneity assumption, $\alpha_\ell \neq 0$, or the exclusion assumption, $\gamma_\ell \neq 0$, but not both, $\gamma_\ell \alpha_\ell = 0$.*

Assumption 5 rules out that a variable considered to be an instrument for the endogenous variable X is itself an endogenous variable.

Note that a valid instrument by itself may not identify β if there are invalid instruments present and when instruments are correlated. A valid instrument then identifies β if the invalid instruments that violate the exclusion restriction are added as controls as specified in model (2) and instruments that violate the exogeneity assumption are removed from the instrument set.

In the next section, we will present the identified set for β, the falsification frontier and the falsification adaptive set as developed by MP for the case where invalid instruments can violate the exclusion assumption only.

The *FAS* is the identified set for β if it contains β. We specify here the condition that the *FAS* is the identified set for β if there is at least one relevant valid instrument.

Condition 1. *Sufficient Condition: For the model specification given in (2) and (3), suppose Assumptions 1, 2 and 5 hold and the joint distribution of $(Y, X, \boldsymbol{Z})$ is known. Then a sufficient condition for the FAS to contain β and thus be the identified set for β is that there is at least one instrument that is relevant and valid.*

Relevance of an instrument will be further clarified below. We consider Condition 1 to be an essential requirement for the *FAS*. We show in Section 3 that this sufficient condition for the FAS_{excl} as proposed by MP requires the validity of the exogeneity Assumption 3 instead of the weaker Assumption 5. In other words, if invalid instruments violate the exogeneity assumption, FAS_{excl} is not guaranteed to contain β. We show this in Section 4, where we develop the FAS_{exo} for the case where instruments may violate the exogeneity Assumption 3 only. The sufficient condition for the FAS_{exo} requires the validity of the exclusion Assumption 4 instead of the weaker Assumption 5. Section 5 then combines these results and proposes the *FAS* for the case that covers the possibility that there are invalid instruments that violate the exclusion restriction and invalid instruments that violate the exogeneity assumption and we show that this *FAS* satisfies Condition 1.

3 Failure of the Exclusion Assumption Only

MP derive the falsification adaptive set for the case of a failure of the exclusion Assumption 4, whilst maintaining the exogeneity Assumption 3. They argue that the same technical analysis can be used to relax both assumptions, something we will discuss in greater detail below.

Let

$$\boldsymbol{\pi} := \boldsymbol{\Sigma}_z^{-1}\text{cov}\left(\boldsymbol{Z}, X\right); \ \boldsymbol{\psi} := \boldsymbol{\Sigma}_z^{-1}\text{cov}\left(\boldsymbol{Z}, Y\right). \tag{4}$$

As it is maintained that $\text{cov}\left(\boldsymbol{Z}, U\right) = \boldsymbol{\alpha} = \mathbf{0}$, it follows from model (2) that $\boldsymbol{\psi} = \boldsymbol{\pi}\beta + \boldsymbol{\gamma}$.

MP make the following partial exclusion assumption.

Assumption 6. *Partial exclusion: There are known constants $\delta_\ell \geq 0$ such that $|\gamma_l| \leq \delta_\ell$ for $\ell = 1, \ldots, k_z$.*

Under Assumptions 1-3 and 6, the identified set for β is then given in MP, Theorem 1, by

$$\mathcal{B}\left(\boldsymbol{\delta}\right) = \left\{b \in \mathbb{R} : -\boldsymbol{\delta} \leq \left(\boldsymbol{\psi} - \boldsymbol{\pi}b\right) \leq \boldsymbol{\delta}\right\}, \tag{5}$$

where the inequalities are componentwise. This follows straightforwardly as $\boldsymbol{\psi} - \boldsymbol{\pi}\beta = \boldsymbol{\gamma}$ and $-\boldsymbol{\delta} \leq \boldsymbol{\gamma} \leq \boldsymbol{\delta}$. The model is falsified if and only if $\mathcal{B}\left(\boldsymbol{\delta}\right)$ is empty. If $\mathcal{B}\left(\mathbf{0}\right)$ is empty then the baseline IV model, assuming $\boldsymbol{\gamma} = \mathbf{0}$, is falsified.

For the individual components of the identified set we have $-\delta_\ell \leq \left(\psi_\ell - \pi_\ell b\right) \leq \delta_\ell$. As stated in Corollary 1 in MP, it follows that

$$\mathcal{B}\left(\boldsymbol{\delta}\right) = \cap_{\ell=1}^{k_z} B_\ell\left(\delta_\ell\right),$$

where

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

The falsification frontier, denoted *FF*, is the minimal set of $\boldsymbol{\delta}$s which lead to a nonempty identified set. For a $\boldsymbol{\delta} \in FF$ this means that for any other $\boldsymbol{\delta}' < \boldsymbol{\delta}$, $B\left(\boldsymbol{\delta}'\right)$ is empty and thus falsifies the model, where $\boldsymbol{\delta}' < \boldsymbol{\delta}$ means that $\delta'_\ell \leq \delta_\ell$ for all $\ell \in \{1, \ldots, k_z\}$ and $\delta'_m < \delta_m$ for some $m \in \{1, \ldots, k_z\}$, see Definition 1 in MP. As in MP, let $\mathcal{L}_{rel}$ denote the set of relevant instruments,

$$\mathcal{L}_{rel} = \{\ell \in \{1, \ldots, k_z\} : \pi_\ell \neq 0\}. \tag{6}$$

For model (2), under Assumptions 1-3 and and 6, Proposition 2 in MP specifies the falsification frontier as the set

$$FF_{excl} = \left\{ \boldsymbol{\delta}(b) \in \mathbb{R}^{k_z}_{\geq 0} : \delta_\ell(b) = |\psi_\ell - b\pi_\ell|, \ \ell = 1, \ldots, k_z, \ b \in \left[\min_{\ell \in \mathcal{L}_{rel}} \frac{\psi_\ell}{\pi_\ell}, \max_{\ell \in \mathcal{L}_{rel}} \frac{\psi_\ell}{\pi_\ell} \right] \right\},$$

where we have added the subscript "excl" to denote a relaxation of the exclusion assumption only, for reasons that will become clear below, and also defined $\boldsymbol{\delta}$ as a function of b.

It is instructive at this point to consider a small numerical example.

Example 1. Let $\beta = 1$, $k_z = 3$, $\boldsymbol{\pi} = (1,1,1)^T$ and $\boldsymbol{\gamma} = (-1,0,2)^T$, resulting in $\boldsymbol{\psi} = (0,1,3)^T$. The values of $\boldsymbol{\delta}(b)$ for values of $b = \{0, 0.5, 1, 3\}$, where $\{0,3\} = \left\{ \min_\ell \frac{\psi_\ell}{\pi_\ell}, \max_\ell \frac{\psi_\ell}{\pi_\ell} \right\}$ are given by

$$\boldsymbol{\delta}(0) = \begin{pmatrix} 0 \\ 1 \\ 3 \end{pmatrix}; \ \boldsymbol{\delta}(0.5) = \begin{pmatrix} 0.5 \\ 0.5 \\ 2.5 \end{pmatrix}; \ \boldsymbol{\delta}(1) = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}; \ \boldsymbol{\delta}(3) = \begin{pmatrix} 3 \\ 2 \\ 0 \end{pmatrix}$$

and for each value of $b \in [0,3]$, it follows that $\mathcal{B}(\boldsymbol{\delta}(b)) = b$. For example, for $\boldsymbol{\delta}(0.5)$,

$$\begin{pmatrix} -0.5 \\ -0.5 \\ -2.5 \end{pmatrix} \leq \left(\begin{pmatrix} 0 \\ 1 \\ 3 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} b \right) \leq \begin{pmatrix} 0.5 \\ 0.5 \\ 2.5 \end{pmatrix}$$

for $b = 0.5$ only. For any $\boldsymbol{\delta}' < \boldsymbol{\delta}(0.5)$ it follows that $\mathcal{B}(\boldsymbol{\delta}')$ is empty. It is clear that for all values of $b \in [0,3]$ it holds that $\mathcal{B}(\boldsymbol{\delta}')$ is empty for any $\boldsymbol{\delta}' < \boldsymbol{\delta}(b)$. For any $b^* \notin [0,3]$ there is a $\boldsymbol{\delta}' < \boldsymbol{\delta}(b^*)$ for which $\mathcal{B}(\boldsymbol{\delta}')$ is not empty, and hence $\boldsymbol{\delta}(b^*) \notin FF$. For example, for $\eta > 0$, $\boldsymbol{\delta}(3+\eta) = (3+\eta, 2+\eta, \eta)^T > \boldsymbol{\delta}(3)$.

The falsification adaptive set, denoted here FAS_{excl}, is then given in Theorem 2 of MP as

$$FAS_{excl} = \cup_{\boldsymbol{\delta} \in FF} \mathcal{B}(\boldsymbol{\delta}) = \left[\min_{\ell \in \mathcal{L}_{rel}} \frac{\psi_\ell}{\pi_\ell}, \max_{\ell \in \mathcal{L}_{rel}} \frac{\psi_\ell}{\pi_\ell} \right]. \tag{7}$$

As MP point out in their Lemma 1, for $\ell \in \mathcal{L}_{rel}$, $\frac{\psi_\ell}{\pi_\ell} \left(= \beta + \frac{\gamma_\ell}{\pi_\ell} \right)$ is the IV/2sls estimand in the just identified model specification

$$Y = X\beta_\ell + \boldsymbol{Z}_{\{-\ell\}}^T \boldsymbol{\gamma}_{\{-\ell\}} + U_\ell \tag{8}$$

where $\boldsymbol{Z}_{\{-\ell\}} = \boldsymbol{Z} \setminus \{Z_\ell\}$, and using Z_ℓ as the excluded just-identifying instrument. This follows directly from the model specification (2) and the linear projection specification $X = \boldsymbol{Z}^T \boldsymbol{\pi} + V$. Hence, $Z_\ell \pi_\ell = X - \boldsymbol{Z}_{\{-\ell\}}^T \boldsymbol{\pi}_{\{-\ell\}} - V$, or $Z_\ell \gamma_\ell = \left(X - \boldsymbol{Z}_{\{-\ell\}}^T \boldsymbol{\pi}_{\{-\ell\}} - V\right) \frac{\gamma_\ell}{\pi_\ell}$. It therefore follows that

$$\beta_\ell = \beta + \frac{\gamma_\ell}{\pi_\ell} = \frac{\psi_\ell}{\pi_\ell};\ \gamma_{\{-\ell\},j} = \gamma_j - \pi_j \frac{\gamma_\ell}{\pi_\ell};\ U_\ell = U - V\frac{\gamma_\ell}{\pi_\ell}$$

and $\text{cov}\left(Z_\ell, U_\ell\right) = 0$. This is for $\ell \in \mathcal{L}_{rel}$, $j = 1, \ldots, k_z, j \neq \ell$, so a slightly unusual notation for $\boldsymbol{\gamma}_{\{-\ell\}}$. For example, for $k_z = 3$ and $\ell = 2$, we have here that $\boldsymbol{\gamma}_{\{-2\}} = \left(\gamma_{\{-2\},1}, \gamma_{\{-2\},3}\right)^T$, see also Windmeijer et al. (2021, Appendix A.5).

As just-identified models are not falsifiable, it follows that for $\ell \in \mathcal{L}_{rel}$, $\delta_\ell\left(\frac{\psi_\ell}{\pi_\ell}\right) = 0$, as highlighted in Example 1. From this, the results of the falsification frontier follow straightforwardly, as when moving b from $\min_{j \in \mathcal{L}_{rel}} \frac{\psi_j}{\pi_j}$ to $\max_{j \in \mathcal{L}_{rel}} \frac{\psi_j}{\pi_j}$ there is at least one element in $\boldsymbol{\delta}(b)$ that decreases in value and at least one that increases in value.

We can write the FAS_{excl} alternatively as

$$FAS_{excl} = \left[\beta + \min_{\ell \in \mathcal{L}_{rel}} \frac{\gamma_\ell}{\pi_\ell}, \beta + \max_{\ell \in \mathcal{L}_{rel}} \frac{\gamma_\ell}{\pi_\ell}\right].$$

It follows that the FAS_{excl} contains β if $0 \in \left[\min_{\ell \in \mathcal{L}_{rel}} \frac{\gamma_\ell}{\pi_\ell}, \max_{\ell \in \mathcal{L}_{rel}} \frac{\gamma_\ell}{\pi_\ell}\right]$ and hence it is then the identified set for β. This is the case in Example 1 where $\beta = 1$ and FAS_{excl} is given by $[0, 3]$. MP, p 1456, state that the FAS_{excl} "...is the identified set for β under the assumption that one of the points on the falsification frontier is true." Maintaining the exogeneity Assumption 3 that $\alpha_\ell = 0$ for all $\ell \in \{1, \ldots, k_z\}$, if there is at least one relevant valid instrument, then $\beta \in FAS_{excl}$, as $\gamma_\ell = 0$ for a valid instrument Z_ℓ. If none of the instruments are valid, the FAS_{excl} contains β if $\min_{\ell \in \mathcal{L}_{rel}} \frac{\gamma_\ell}{\pi_\ell} < 0$ and $\max_{\ell \in \mathcal{L}_{rel}} \frac{\gamma_\ell}{\pi_\ell} > 0$. Therefore, under the exogeneity assumption, a sufficient, but not necessary condition for FAS_{excl} to be the identified set for β is that at least one of the instruments is valid and relevant. We state this in the following condition.

Condition 2. *Exclusion Sufficient Condition: If the instruments satisfy the exogeneity Assumption 3 that $\alpha_\ell = 0$ for all $\ell \in \{1, \ldots, k_z\}$, then a sufficient condition for the FAS_{excl} to be the identified set for β is that there is at least one instrument that is relevant and* *valid.*

As specified for $\mathcal{L}_{rel}$ in (6), an instrument is relevant in this setting if $\pi_\ell \neq 0$. Condition 2 requires the exogeneity Assumption 3 which is a stronger assumption than Assumption 5 that an invalid instrument can violate the exogeneity or exclusion assumption for the general sufficient Condition 1.

3.1 Related Instrument Selection Methods

Kang et al. (2016), Guo et al. (2018), Windmeijer et al. (2019) and Windmeijer et al. (2021) considered the same setting where invalid instruments can only violate the exclusion assumption. These papers then salvage falsified models by selecting the set of valid instruments under either a majority rule or plurality rule. The majority rule states that more than 50% of the relevant instruments are valid, whilst the plurality rule states that within $\mathcal{L}_{rel}$ the valid instruments form the largest group, where a group of instruments is defined by having the same value $\frac{\psi_\ell}{\pi_\ell}$.

The latter can be seen in relation to the falsification frontier, by finding, within $\mathcal{L}_{rel}$, the value of $b = \frac{\psi_\ell}{\pi_\ell}$ for which $\boldsymbol{\delta}(b)$ has the largest number of zeros. To illustrate, if we add a fourth instrument to our numerical Example 1 above, with $\pi_4 = 1$ and $\gamma_4 = 0$, then we have both Z_2 and Z_4 as valid instruments and they form the largest group, as the values of $\frac{\psi_\ell}{\pi_\ell}$ are given by $\{0, 1, 3, 1\}$. For $b = \beta = 1$ we then clearly have that $\boldsymbol{\delta}(1) = (1, 0, 2, 0)^T$ and the selection methods of Guo et al. (2018) and Windmeijer et al. (2021) consistently select Z_2 and Z_4 as the valid instruments. These two selection methods use dimension reduction techniques to deal with situations of a large number of instruments. For a small number of instruments, as in this example, the selection procedures of Andrews (1999), adjusted to deal with the violation of the exclusion assumption only, can be used to select the set of valid instruments.

3.2 Estimation of $\boldsymbol{FAS}_{excl}$ and First-Stage Hard Thresholding

We have an i.i.d. sample of size n, $\left\{Y_i, X_i, \boldsymbol{Z}_i^T\right\}_{i=1}^n$. The n-vectors (Y_i) and (X_i) are denoted $\boldsymbol{y}$ and $\boldsymbol{x}$ respectively, and here $\boldsymbol{Z}$ denotes the $k_z \times n$ matrix of observations on the instruments. MP suggest to estimate the set of relevant instruments by

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

where F_ℓ is the first-stage F-statistic for model (8), where Z_ℓ is considered as an instrument and $Z_{\{-\ell\}}$ as controls. For all values of ℓ, the first-stage model is therefore given by $\boldsymbol{x} = \boldsymbol{Z}\boldsymbol{\pi} + \boldsymbol{v}$ and so F_ℓ is the same as the Wald statistic for testing $H_0 : \pi_\ell = 0$ based on the OLS estimator of $\boldsymbol{\pi}$, denoted $\widehat{\boldsymbol{\pi}}$. The same first-stage hard thresholding was proposed in Guo et al. (2018). Although $C_n \to \infty$ as $n \to \infty$ and $C_n = o(n)$ for consistent selection, MP choose $C_n = 10$ as their default cutoff, or for the t-ratio, $\left|\frac{\widehat{\pi}_\ell}{se(\widehat{\pi}_\ell)}\right| \geq \sqrt{10} = 3.16$, where $\widehat{\pi}_\ell$ is the OLS estimator of π_ℓ in the first-stage model.

Let $\widehat{\beta}_\ell$ be the IV estimator of β_ℓ in just-identified model specification (8). Then FAS_{excl} is estimated by

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

and MP show that $\widehat{FAS}_{excl}$ is a consistent estimator of the FAS_{excl} under the conditions of their Proposition 3.

Note that instead of having to calculate $k_{z,rel} = \left|\widehat{\mathcal{L}}_{rel}\right|$ IV estimates, the same estimates are obtained from two OLS regressions. Let $\widehat{\boldsymbol{\psi}}$ denote the OLS estimator of $\boldsymbol{\psi}$ in the reduced form model $\boldsymbol{y} = \boldsymbol{Z}\boldsymbol{\psi} + \boldsymbol{v}_y$. Let $\boldsymbol{z}_\ell$ denote the n-vector of observations on the ℓ-th instrument $(Z_{i,\ell})$, the ℓ-th column of $\boldsymbol{Z}$, and

$$\boldsymbol{z}_{\ell|\{-\ell\}} := \boldsymbol{M}_{Z_{\{-\ell\}}} \boldsymbol{z}_\ell,$$

the residual vector after regressing $\boldsymbol{z}_\ell$ on $\boldsymbol{Z}_{\{-\ell\}}$, where $\boldsymbol{Z}_{\{-\ell\}}$ is the $n \times (k_z - 1)$ matrix $\boldsymbol{Z} \setminus \boldsymbol{z}_\ell$, and where for a general full column rank matrix $\boldsymbol{A}$, $\boldsymbol{M}_A = \boldsymbol{I}_n - \boldsymbol{P}_A$, with $\boldsymbol{P}_A = \boldsymbol{A}\left(\boldsymbol{A}^T\boldsymbol{A}\right)^{-1}\boldsymbol{A}^T$ and $\boldsymbol{I}_n$ is the n-dimensional identity matrix. Then it follows that

$$\widehat{\beta}_\ell = \frac{\boldsymbol{z}_{\ell|\{-\ell\}}^T \boldsymbol{y}}{\boldsymbol{z}_{\ell|\{-\ell\}}^T \boldsymbol{x}} = \frac{\left(\boldsymbol{z}_{\ell|\{-\ell\}}^T \boldsymbol{z}_{\ell|\{-\ell\}}\right)^{-1} \boldsymbol{z}_{\ell|\{-\ell\}}^T \boldsymbol{y}}{\left(\boldsymbol{z}_{\ell|\{-\ell\}}^T \boldsymbol{z}_{\ell|\{-\ell\}}\right)^{-1} \boldsymbol{z}_{\ell|\{-\ell\}}^T \boldsymbol{x}} = \frac{\widehat{\psi}_\ell}{\widehat{\pi}_\ell}, \tag{9}$$

see also Windmeijer et al. (2021, Appendix A.5). Note further that irrelevant instruments $\boldsymbol{z}_j$ for $j \notin \widehat{\mathcal{L}}_{rel}$ are included as control variables in this setting, as it is in Guo et al. (2018) and Windmeijer et al. (2021).

3.3 2sls

For the model specification

$$\boldsymbol{y} = \boldsymbol{x}\beta + \boldsymbol{u},$$

and the $n \times k_z$ instrument matrix $\boldsymbol{Z}$, the 2sls estimator is given by

$$\begin{aligned}\widehat{\beta}_{2sls} &= \left(\boldsymbol{x}^T \boldsymbol{P}_Z \boldsymbol{x}\right)^{-1} \boldsymbol{x}^T \boldsymbol{P}_Z \boldsymbol{y} \\ &= \left(\boldsymbol{x}^T \boldsymbol{P}_Z \boldsymbol{x}\right)^{-1} \boldsymbol{x}^T \boldsymbol{Z} \widehat{\boldsymbol{\psi}} \\ &= \left(\boldsymbol{x}^T \boldsymbol{P}_Z \boldsymbol{x}\right)^{-1} \boldsymbol{x}^T \boldsymbol{Z} \boldsymbol{D}_{\widehat{\pi}} \widehat{\boldsymbol{\beta}}\end{aligned}$$

where $\widehat{\boldsymbol{\beta}}$ is the k_z-vector $\left(\widehat{\beta}_\ell\right)$, and $\boldsymbol{D}_{\widehat{\pi}}$ is a $k_z \times k_z$ diagonal matrix with ℓ-th diagonal element equal to $\widehat{\pi}_\ell$. Therefore $\widehat{\beta}_{2sls}$ is a linear combination of the just-identified IV estimators,

$$\widehat{\beta}_{2sls} = \sum_{\ell=1}^{k_z} w_\ell \widehat{\beta}_\ell, \tag{10}$$

with $w_\ell = \widehat{\pi}_\ell \boldsymbol{z}_\ell^T \boldsymbol{x} \left(\boldsymbol{x}^T \boldsymbol{P}_Z \boldsymbol{x}\right)^{-1}$. Let $\boldsymbol{\iota}_{k_z}$be a k_z-vector of ones. As $\boldsymbol{x}^T \boldsymbol{Z} \boldsymbol{D}_{\widehat{\pi}} \boldsymbol{\iota}_{k_z} = \boldsymbol{x}^T \boldsymbol{Z} \widehat{\boldsymbol{\pi}} = \boldsymbol{x}^T \boldsymbol{P}_Z \boldsymbol{x}$ it follows that $\sum_{\ell=1}^{k_z} w_\ell = 1$, but $\widehat{\beta}_{2sls}$ is not necessarily a weighted average of the $\widehat{\beta}_\ell$ as w_ℓ can be negative. This is the case when $\widehat{\pi}_\ell$ and $\boldsymbol{z}_\ell^T \boldsymbol{x}$ have opposite signs, or equivalently if $\widehat{\pi}_\ell$ and $\widehat{\pi}_\ell^*$ have opposite signs, where $\widehat{\pi}_\ell^*$ is the OLS estimator of π_ℓ^*, the coefficient in the linear specification $\boldsymbol{x} = \boldsymbol{z}_\ell \pi_\ell^* + \boldsymbol{v}_\ell$.

4 Failure of the Exogeneity Assumption Only

We now maintain the exclusion Assumption 4, so $\boldsymbol{\gamma} = \boldsymbol{0}$, but consider violations of the exogeneity Assumption 3. The observed outcome model is therefore given by

$$Y = X\beta + U, \tag{11}$$

with the possible exogeneity violation specified as in (3),

$$\operatorname{cov}(\boldsymbol{Z}, U) = \boldsymbol{\alpha}.$$

MP, page 1453, argue that, mathematically, the same technical analysis can be used to relax the exogeneity assumption as is used above for the relaxation of the exclusion assumption. In MP20, Appendix G, it is argued that a linear projection of U on $\boldsymbol{Z}$ results in

$$U = \boldsymbol{Z}^T \boldsymbol{\eta} + \dot{U}, \tag{12}$$

with $\text{cov}\left(\boldsymbol{Z}, \dot{U}\right) = \mathbf{0}$ by construction. Then

$$\begin{aligned} Y &= X\beta + U \\ &= X\beta + \boldsymbol{Z}^T\boldsymbol{\eta} + \dot{U}, \end{aligned}$$

which is the same type of specification as that of model (2) and MP20 argue that the results above for the relaxation of the exclusion assumption apply here as well. MP20 stress that the key difference is the interpretation of the coefficients on $\boldsymbol{Z}$. We argue here that it actually has an implication for the *FAS*.

For the linear projection specification (12), we have that

$$\boldsymbol{\eta} = \boldsymbol{\Sigma}_z^{-1}\text{cov}\left(\boldsymbol{Z}, U\right) = \boldsymbol{\Sigma}_z^{-1}\boldsymbol{\alpha}. \tag{13}$$

If the instruments are orthogonal or independent and $\boldsymbol{\Sigma}_z$ is a diagonal matrix, then the distinction between a relaxation of the exclusion or exogeneity assumption is immaterial. However, for a general $\boldsymbol{\Sigma}_z$, the two different relaxations need different treatment. With correlated instruments, instruments that violate the exclusion restriction will need to be included as controls in the model for a valid instrument to identify β. Instruments that violate the exogeneity assumption need to be excluded from the instrument set for a valid instrument to identify β.

As stated in Condition 2, when considering a relaxation of the exclusion assumption only, a sufficient condition for the FAS_{excl} as defined in (7) to contain β and hence to be the identified set for β is that at least one instrument is relevant and valid, satisfying the exclusion assumption $\gamma_\ell = 0$. For a relaxation of the exogeneity assumption only, there could be a relevant and valid instrument with $\alpha_\ell = 0$, but from (13) it follows that it could be the case that $\eta_\ell \neq 0$ for $\ell = 1, \ldots, k_z$, due to the correlation of the instruments. So although there is a valid instrument, the FAS_{excl} of (7) is then not guaranteed to contain β. A *FAS* that does is easily obtained.

Like the partial exclusion Assumption 6 as made in MP, we make the following partial exogeneity assumption.

Assumption 7. *Partial exogeneity: There are known constants* $\delta_\ell^* \geq 0$ *such that* $|\alpha_\ell^*| \leq \delta_\ell^*$ *for* $\ell = 1, \ldots, k_z$*, where* $\alpha_\ell^* = \left(var\left(Z_\ell\right)\right)^{-1} \alpha_\ell$*.*

Then define the k_z-vectors $\boldsymbol{\pi}^*$ and $\boldsymbol{\psi}^*$ with ℓ-th elements given by

$$\pi_\ell^* := (\mathrm{var}\,(Z_\ell))^{-1}\,\mathrm{cov}\,(Z_\ell, X)\,;\;\; \psi_\ell^* := (\mathrm{var}\,(Z_\ell))^{-1}\,\mathrm{cov}\,(Z_\ell, Y)\,, \tag{14}$$

for $\ell = 1, \ldots, k_z$.

Under Assumptions 1, 2, 4 and 7, the identified set for β is given by

$$\mathcal{B}\,(\boldsymbol{\delta}^*) = \{b \in \mathbb{R} : -\boldsymbol{\delta}^* \leq (\boldsymbol{\psi}^* - \boldsymbol{\pi}^* b) \leq \boldsymbol{\delta}^*\}\,,$$

which follows straightforwardly as $\boldsymbol{\psi}^* - \boldsymbol{\pi}^*\beta = \boldsymbol{\alpha}^*$ and $-\boldsymbol{\delta}^* \leq \boldsymbol{\alpha}^* \leq \boldsymbol{\delta}^*$. The model is falsified if and only if $\mathcal{B}\,(\boldsymbol{\delta}^*)$ is empty. If $\mathcal{B}\,(\mathbf{0})$ is empty then the baseline IV model (11), assuming $\boldsymbol{\alpha} = \mathbf{0}$, is falsified.

Let $\mathcal{L}_{rel}^*$ denote the set of relevant instruments

$$\mathcal{L}_{rel}^* = \{\ell \in \{1, \ldots, k_z\} : \pi_\ell^* \neq 0\}\,. \tag{15}$$

For the IV model (11), under Assumptions 1, 2 and 4, known joint distribution of $(Y, X, \boldsymbol{Z})$ and following the same arguments as above for the derivation of the falsification frontier for the relaxation of exclusion restrictions, the falsification frontier here is given as the set

$$FF_{exo} = \left\{\boldsymbol{\delta}^*(b) \in \mathbb{R}_{\geq 0}^{k_z} : \delta_\ell^*(b) = |\psi_\ell^* - b\pi_\ell^*|\,,\; \ell = 1, \ldots, k_z,\; b \in \left[\min_{\ell \in \mathcal{L}_{rel}^*} \frac{\psi_\ell^*}{\pi_\ell^*}, \max_{\ell \in \mathcal{L}_{rel}^*} \frac{\psi_\ell^*}{\pi_\ell^*}\right]\right\}.$$

The falsification adaptive set is therefore given

$$\begin{aligned} FAS_{exo} &= \left[\min_{\ell \in \mathcal{L}_{rel}^*} \frac{\psi_\ell^*}{\pi_\ell^*}, \max_{\ell \in \mathcal{L}_{rel}^*} \frac{\psi_\ell^*}{\pi_\ell^*}\right] \\ &= \left[\beta + \min_{\ell \in \mathcal{L}_{rel}^*} \frac{\alpha_\ell^*}{\pi_\ell^*}, \beta + \max_{\ell \in \mathcal{L}_{rel}^*} \frac{\alpha_\ell^*}{\pi_\ell^*}\right]. \end{aligned} \tag{16}$$

For $\ell \in \mathcal{L}_{rel}^*$, the ratio

$$\beta_\ell^* := \frac{\psi_\ell^*}{\pi_\ell^*} = \frac{\mathrm{cov}\,(Z_\ell, Y)}{\mathrm{cov}\,(Z_\ell, X)}$$

is the IV estimand for the specification $Y = X\beta + U$, using Z_ℓ as the just-identifying instrument for X, and treating the instruments $\boldsymbol{Z}_{\{-\ell\}}$ as invalid and excluding them from the analysis.

Under the exclusion assumption, a sufficient condition for FAS_{exo} to contain β is that there is at least one relevant and valid instrument as we state in the following condition.

Condition 3. *Exogeneity Sufficient Condition: If the instruments satisfy the exclusion*

Assumption 4 that $\gamma_\ell = 0$ for all $\ell \in \{1, \dots, k_z\}$, then a sufficient condition for the FAS_{exo} to be the identified set for β is that at least one instrument is relevant and valid.

As specified for $\mathcal{L}^*_{rel}$ in (15), an instrument Z_ℓ is relevant in this setting if $\pi^*_\ell \neq 0$. Relative to Assumption 5 for the general Condition 1, Condition 3 requires the stronger exclusion assumption.

To contrast the condition for FAS_{exo} with the finding for FAS_{excl}, consider the following simple example where there is one valid instrument and one invalid instrument that violates the exogeneity assumption. As Condition 2 is not satisfied, the FAS_{excl} is in this case not guaranteed to be the identified set for β, even though one of the instruments is valid and relevant.

Example 2. Let $k_z = 2$, $\boldsymbol{\pi} = (\pi_1, \pi_2)^T$, $\pi_\ell \neq 0$ for $\ell = 1, 2$, $\boldsymbol{\gamma} = \mathbf{0}$, $\boldsymbol{\alpha} = (0, \alpha_2)^T$, $\alpha_2 \neq 0$, and

$$\boldsymbol{\Sigma}_z = \begin{bmatrix} 1 & \rho_{12} \\ \rho_{12} & 1 \end{bmatrix},$$

with $|\rho_{12}| < 1$. Then $\boldsymbol{\pi}^* = (\pi_1 + \rho_{12}\pi_2, \pi_2 + \rho_{12}\pi_1)^T$, $\boldsymbol{\alpha}^* = \boldsymbol{\alpha}$, and so, provided $\pi^*_\ell \neq 0$ for $\ell = 1, 2$,

$$FAS_{exo} = \beta + \left[\min\left\{0, \frac{\alpha_2}{\pi_2 + \rho_{12}\pi_1}\right\}, \max\left\{0, \frac{\alpha_2}{\pi_2 + \rho_{12}\pi_1}\right\}\right]$$

and $\beta \in FAS_{exo}$.

For FAS_{excl}, we have that $\boldsymbol{\eta} = \boldsymbol{\Sigma}_z^{-1}\boldsymbol{\alpha} = \frac{\alpha_2}{1-\rho_{12}^2}(-\rho_{12}, 1)^T$, and so

$$FAS_{excl} = \beta + \left[\min\left\{\frac{-\rho_{12}\alpha_2}{\pi_1(1-\rho_{12}^2)}, \frac{\alpha_2}{\pi_2(1-\rho_{12}^2)}\right\}, \max\left\{\frac{-\rho_{12}\alpha_2}{\pi_1(1-\rho_{12}^2)}, \frac{\alpha_2}{\pi_2(1-\rho_{12}^2)}\right\}\right],$$

therefore, if $\frac{-\rho_{12}\alpha_2}{\pi_1}$ and $\frac{\alpha_2}{\pi_2}$ have the same sign, $\beta \notin FAS_{excl}$.

We can of course adapt Example 2 to the case where the invalid instrument violates the exclusion assumption instead, in which case the FAS_{exo} is not guaranteed to be the identified set for β.

The selection methods of Guo et al. (2018) and Windmeijer et al. (2021) can be adjusted to select the valid instruments consistently in the case of a violation of the exogeneity assumption only if the plurality rule applies that the valid instruments form the largest group within the set of relevant variables. The Andrews (1999) selection method was originally designed for this case.

4.1 Estimation

The set of relevant instruments can here be estimated by

$$\widehat{\mathcal{L}}_{rel}^{*}=\left\{\ell\in\{1,\ldots,k_z\}:F_\ell^{*}\geq C_n\right\},$$

where F_ℓ^{*} is the F-statistic for testing $H_0:\pi_\ell^{*}=0$ in the first-stage linear specification

$$\boldsymbol{x}=\boldsymbol{z}_\ell\pi_\ell^{*}+\boldsymbol{v}_\ell.$$

Let $\widehat{\beta}_\ell^{*}$ be the IV estimators

$$\widehat{\beta}_\ell^{*}=\frac{\boldsymbol{z}_\ell^T\boldsymbol{y}}{\boldsymbol{z}_\ell^T\boldsymbol{x}}=\frac{\widehat{\psi}_\ell^{*}}{\widehat{\pi}_\ell^{*}},$$

where $\widehat{\psi}_\ell^{*}$ is the OLS estimator of ψ_ℓ^{*} in the reduced-form specification $\boldsymbol{y}=\boldsymbol{z}_\ell^T\psi_\ell^{*}+\boldsymbol{\varepsilon}_\ell$, and $\widehat{\pi}_\ell^{*}$ the OLS estimator of π_ℓ^{*}. Then the consistent estimator of FAS_{exo} is given by

$$\widehat{FAS}_{exo}=\left[\min_{\ell\in\widehat{\mathcal{L}}_{rel}^{*}}\widehat{\beta}_\ell^{*},\max_{\ell\in\widehat{\mathcal{L}}_{rel}^{*}}\widehat{\beta}_\ell^{*}\right].$$

4.2 2sls

The 2sls estimator for β for the specification $\boldsymbol{y}=\boldsymbol{x}\beta+\boldsymbol{u}$ using the $n\times k_z$ instrument matrix $\boldsymbol{Z}$ is again a linear combination of the individual estimators $\widehat{\beta}_\ell^{*}$,

$$\begin{aligned}\widehat{\beta}_{2sls}&=\left(\boldsymbol{x}^T\boldsymbol{P}_Z\boldsymbol{x}\right)^{-1}\boldsymbol{x}^T\boldsymbol{P}_Z\boldsymbol{y}\\&=\left(\boldsymbol{x}^T\boldsymbol{P}_Z\boldsymbol{x}\right)^{-1}\widehat{\boldsymbol{\pi}}^T\boldsymbol{D}_{Z^Tx}\widehat{\boldsymbol{\beta}}^{*}\\&=\sum_{\ell=1}^{k_z}w_\ell^{*}\widehat{\beta}_\ell^{*}.\end{aligned}\tag{17}$$

where $\widehat{\boldsymbol{\beta}}^{*}$ is the k_z-vector $\left(\widehat{\beta}_\ell^{*}\right)$ and $\boldsymbol{D}_{Z^Tx}$ is the $k_z\times k_z$ diagonal matrix with ℓ-th diagonal element equal to $\boldsymbol{z}_\ell^T\boldsymbol{x}$. It follows that $\left(\boldsymbol{x}^T\boldsymbol{P}_Z\boldsymbol{x}\right)^{-1}\widehat{\boldsymbol{\pi}}^T\boldsymbol{D}_{Z^Tx}\boldsymbol{\iota}_{\boldsymbol{k_z}}=1$ and so $\sum_{\ell=1}^{k_z}w_\ell^{*}=1$. Comparing the weights with those in (10), where we had that $\widehat{\beta}_{2sls}=\sum_{\ell=1}^{k_z}w_\ell\widehat{\beta}_\ell$, with $\sum_{\ell=1}^{k_z}w_\ell=1$, we find that the weights are identical, as

$$w_\ell^{*}=\widehat{\pi}_\ell\boldsymbol{z}_\ell^T\boldsymbol{x}\left(\boldsymbol{x}^T\boldsymbol{P}_Z\boldsymbol{x}\right)^{-1}=w_\ell,$$

for $\ell=1,\ldots,k_z$, i.e. the 2sls weights for the just-identified estimators $\widehat{\beta}_\ell$, where $\boldsymbol{Z}_{\{-\ell\}}$ are included as controls, are the same as those for the just-identified estimators $\widehat{\beta}_\ell^{*}$, where

$\boldsymbol{Z}_{\{-\ell\}}$ are excluded from the instrument set.

5 Failure of the Exogeneity or Exclusion Assumption

If a model is falsified, a researcher is unlikely to know whether this is due to a violation of the exogeneity or the exclusion assumption and the falsification adaptive set should reflect this. In this section we generalize the *FAS*, such that it reflects the fact that instruments can violate the exogeneity or exclusion restriction. We therefore consider here the general model specifications as in (2) and (3),

$$Y = X\beta + \boldsymbol{Z}^T\boldsymbol{\gamma} + U$$

$$\operatorname{cov}(\boldsymbol{Z}, U) = \boldsymbol{\alpha}.$$

We obtain the falsification frontier and associated *FAS* that satisfies Condition 1 by simply considering all possible just-identified model specifications, as detailed below. We first show this for the $k_z = 2$ case, where the results are a simple combination of the results for the violation of the exclusion restriction only and the violation of the exogeneity assumption only.

5.1 $k_z = 2$

When there are two instruments available, we can generalize the *FAS* by simply combining the results for FAS_{excl} and FAS_{exo} as presented in Sections 3 and 4. There are in this case four just-identified model specifications, one with Z_1 as control and Z_2 as excluded instrument for X, one with Z_2 as control and Z_1 as excluded instrument for X, and the model without controls and Z_1 or Z_2 as instrument for X. As above, let $\boldsymbol{\pi}$ and $\boldsymbol{\psi}$ be as defined in (4) for the model with instruments included as controls, and $\boldsymbol{\pi}^*$ and $\boldsymbol{\psi}^*$ for the model without instruments included as controls, as defined in (14).

From the combination of the results in Sections 3 and 4, it follows that we get here

$$\boldsymbol{\psi} - \boldsymbol{\pi}\beta = \boldsymbol{\gamma} + \boldsymbol{\eta},$$

with, as in (13), $\boldsymbol{\eta} = \boldsymbol{\Sigma}_z^{-1}\boldsymbol{\alpha}$. Further

$$\boldsymbol{\psi}^* - \boldsymbol{\pi}^*\beta = \boldsymbol{\alpha}^* + \boldsymbol{\xi}^*,$$

with $\boldsymbol{\xi}^* = \text{diag}\left((\text{var}(Z_\ell))^{-1}\right)\boldsymbol{\Sigma}_z\boldsymbol{\gamma}$, where for general d_ℓ, diag(d_ℓ) is a diagonal matrix with ℓ-th diagonal element d_ℓ, and, as before, $\alpha^*_\ell = (\text{var}(Z_\ell))^{-1}\alpha_\ell$. We can now make the partial exogeneity and exclusion assumption.

Assumption 8. *Partial exogeneity and exclusion: There are known constants $\delta_\ell \geq 0$ and $\delta^*_\ell \geq 0$ such that $|\gamma_\ell + \eta_\ell| \leq \delta_\ell$ and $|\alpha^*_\ell + \xi^*_\ell| \leq \delta^*_\ell$ for all $\ell \in \{1, \ldots, k_z\}$.*

Then define

$$\tilde{\boldsymbol{\pi}} := \begin{pmatrix} \boldsymbol{\pi} \\ \boldsymbol{\pi}^* \end{pmatrix};\ \tilde{\boldsymbol{\psi}} := \begin{pmatrix} \boldsymbol{\psi} \\ \boldsymbol{\psi}^* \end{pmatrix};\ \tilde{\boldsymbol{\delta}} := \begin{pmatrix} \boldsymbol{\delta} \\ \boldsymbol{\delta}^* \end{pmatrix}.$$

Extending the results as outlined in Sections 3 and 4, under Assumptions 1, 2, 5 and 8, the identified set for β is given by

$$\mathcal{B}\left(\tilde{\boldsymbol{\delta}}\right) = \left\{b \in \mathbb{R} : -\tilde{\boldsymbol{\delta}} \leq \left(\tilde{\boldsymbol{\psi}} - \tilde{\boldsymbol{\pi}}b\right) \leq \tilde{\boldsymbol{\delta}}\right\}.$$

Let $\tilde{\mathcal{L}}_{rel}$ denote the set of relevant instruments in this setting, specified as

$$\tilde{\mathcal{L}}_{rel} = \{j \in \{1, \ldots, 4\} : \tilde{\pi}_j \neq 0\}. \tag{18}$$

Then we get for the falsification frontier,

$$FF = \left\{\tilde{\boldsymbol{\delta}}(b) \in \mathbb{R}^4_{\geq 0} : \tilde{\delta}_j(b) = \left|\tilde{\psi}_j - b\tilde{\pi}_j\right|,\ j = 1, \ldots, 4,\ b \in \left[\min_{j \in \tilde{\mathcal{L}}_{rel}} \frac{\tilde{\psi}_j}{\tilde{\pi}_j}, \max_{j \in \tilde{\mathcal{L}}_{rel}} \frac{\tilde{\psi}_j}{\tilde{\pi}_j}\right]\right\}.$$

As discussed in Section 3, this follows as, for $j \in \tilde{\mathcal{L}}_{rel}$, $\tilde{\delta}_j\left(\frac{\tilde{\psi}_j}{\tilde{\pi}_j}\right) = 0$ and so when moving b from $\min_{j \in \tilde{\mathcal{L}}_{rel}} \frac{\tilde{\psi}_j}{\tilde{\pi}_j}$ to $\max_{j \in \tilde{\mathcal{L}}_{rel}} \frac{\tilde{\psi}_j}{\tilde{\pi}_j}$ there is at least one element in $\tilde{\boldsymbol{\delta}}(b)$ that decreases in value and at least one that increases in value. For any $\tilde{\boldsymbol{\delta}}' < \tilde{\boldsymbol{\delta}}(b)$, $b \in \left[\min_{j \in \tilde{\mathcal{L}}_{rel}} \frac{\tilde{\psi}_j}{\tilde{\pi}_j}, \max_{j \in \tilde{\mathcal{L}}_{rel}} \frac{\tilde{\psi}_j}{\tilde{\pi}_j}\right]$, $\mathcal{B}\left(\tilde{\boldsymbol{\delta}}'\right)$ is empty.

We then get our main result that the falsification frontier is given by

$$FAS = \left[\min_{j \in \tilde{\mathcal{L}}_{rel}} \frac{\tilde{\psi}_j}{\tilde{\pi}_j}, \max_{j \in \tilde{\mathcal{L}}_{rel}} \frac{\tilde{\psi}_j}{\tilde{\pi}_j}\right]. \tag{19}$$

These results are clear. In order to consider possible violations of the exogeneity or exclusion assumptions, we need to consider all just-identified model specifications.

When there are two instruments available, the model is falsified and one instrument is valid, then from Assumption 5 it follows that the invalid instrument does either violate the exogeneity or the exclusion restriction. If it violates the exogeneity assumption, as in Example 2, with $\boldsymbol{\alpha} = (0, \alpha_2)^T$, $\alpha_2 \neq 0$, then Z_1 identifies β provided Z_2 is omitted from the instrument set as $\frac{\tilde{\psi}_3}{\tilde{\pi}_3} = \frac{\psi_1^*}{\pi_1^*} = \beta$, provided Z_1 is a relevant instrument when Z_2 is omitted, or $\tilde{\pi}_3 = \pi_1^* \neq 0$. If alternatively the invalid instrument violates the exclusion assumption, with $\boldsymbol{\gamma} = (0, \gamma_2)^T$, $\gamma_2 \neq 0$, then Z_1 identifies β when Z_2 is included as a control variable as $\frac{\tilde{\psi}_1}{\tilde{\pi}_1} = \frac{\psi_1}{\pi_1} = \beta$, provided Z_1 is a relevant instrument when Z_2 is included as a control and hence in the instrument set, or $\tilde{\pi}_1 = \pi_1 \neq 0$. In the words of MP, p 1452, when the baseline model is falsified, the falsification adaptive set expands to account for the uncertainty about which assumption along the frontier is true. The *FAS* as defined in (19) does this for violations of the exogeneity and exclusion assumptions. As it contains β if there is a valid relevant instrument, it satisfies Condition 1.

5.1.1 Alternative Representation

Instead of specifying different just identified models as in the previous sections, we can simply focus on the transformed just-identifying instruments, which will enable us to easily generalize the approach to the setup for a general number of instruments. Define the population linear projection errors $Z_{1|2}$ and $Z_{2|1}$ as

$$Z_{1|2} := Z_1 - Z_2\phi_{21};\ Z_{2|1} := Z_2 - Z_1\phi_{12}, \tag{20}$$

where $\phi_{21} = (\text{var}(Z_2))^{-1} \text{cov}(Z_2, Z_1)$ and $\phi_{12} = (\text{var}(Z_1))^{-1} \text{cov}(Z_1, Z_2)$. It follows straightforwardly that for the elements of $\boldsymbol{\psi} = \boldsymbol{\Sigma}_z^{-1}\text{cov}(\boldsymbol{Z}, Y)$, we have that

$$\psi_1 = \left(\text{var}\left(Z_{1|2}\right)\right)^{-1} \text{cov}\left(Z_{1|2}, Y\right)$$
$$\psi_2 = \left(\text{var}\left(Z_{2|1}\right)\right)^{-1} \text{cov}\left(Z_{2|1}, Y\right),$$

and equivalently for the elements of $\boldsymbol{\pi}$. These are the population equivalents of the sample estimation results in (9).

We can therefore get the vector of just-identifying transformed instruments as

$$\tilde{\boldsymbol{Z}} = \left(Z_{1|2}, Z_{2|1}, Z_1, Z_2\right)^T. \tag{21}$$

Then the vectors $\tilde{\boldsymbol{\pi}}$ and $\tilde{\boldsymbol{\psi}}$ have the elements

$$\tilde{\pi}_j = \left(\text{var}\left(\tilde{Z}_j\right)\right)^{-1}\text{cov}\left(\tilde{Z}_j, X\right);\ \tilde{\psi}_j = \left(\text{var}\left(\tilde{Z}_j\right)\right)^{-1}\text{cov}\left(\tilde{Z}_j, Y\right), \tag{22}$$

for $j = 1, \ldots, 4$.

For $j \in \tilde{\mathcal{L}}_{rel}$, the IV estimands are then

$$\tilde{\beta}_j = \frac{\tilde{\psi}_j}{\tilde{\pi}_j} = \frac{\text{cov}\left(\tilde{Z}_j, Y\right)}{\text{cov}\left(\tilde{Z}_j, X\right)} = \beta + \frac{\text{cov}\left(\tilde{Z}_j, \tilde{U}\right)}{\text{cov}\left(\tilde{Z}_j, X\right)},$$

where $\tilde{U} = Y - X\beta = \boldsymbol{Z}^T\boldsymbol{\gamma} + U$. The FAS is given by

$$FAS = \left[\min_{j \in \tilde{\mathcal{L}}_{rel}} \tilde{\beta}_j, \max_{j \in \tilde{\mathcal{L}}_{rel}} \tilde{\beta}_j\right].$$

This representation makes the definition of a relevant instrument generic, with an instrument $\tilde{Z}_j$ here relevant if $\tilde{\pi}_j \neq 0$, confirming the definition as in (18).

5.1.2 Estimation

For estimation of the FAS as defined in (19) the set of relevant instruments is estimated by

$$\widehat{\tilde{\mathcal{L}}}_{rel} = \left\{j \in \{1, \ldots, 4\} : \tilde{F}_j \geq C_n\right\},$$

where $\tilde{F}_j$ is the F-statistic for testing $H_0 : \tilde{\pi}_j = 0$ in the first-stage linear specification $\boldsymbol{x} = \tilde{\boldsymbol{z}}_j\tilde{\pi}_j + \boldsymbol{v}_j$, where $\tilde{\boldsymbol{z}}_j$ is the j-th column of the $n \times 4$ matrix of just-identifying transformed instruments

$$\tilde{\boldsymbol{Z}} = \left[\boldsymbol{M}_{z_2}\boldsymbol{z}_1\ \boldsymbol{M}_{z_1}\boldsymbol{z}_2\ \boldsymbol{z}_1\ \boldsymbol{z}_2\right].$$

Let $\widehat{\tilde{\beta}}_j$ be the IV estimators

$$\widehat{\tilde{\beta}}_j = \frac{\tilde{\boldsymbol{z}}_j^T\boldsymbol{y}}{\tilde{\boldsymbol{z}}_j^T\boldsymbol{x}} = \frac{\widehat{\tilde{\psi}}_j}{\widehat{\tilde{\pi}}_j},$$

where $\widehat{\tilde{\psi}}_j$ is the OLS estimator of $\tilde{\psi}_j$ in the reduced-form specification $\boldsymbol{y} = \tilde{\boldsymbol{z}}_j^T\psi_j + \boldsymbol{\varepsilon}_j$, and $\widehat{\tilde{\pi}}_j$ the OLS estimator of $\tilde{\pi}_j$. Then a consistent estimator of FAS is given by

$$\widehat{FAS} = \left[\min_{j \in \widehat{\tilde{\mathcal{L}}}_{rel}} \widehat{\tilde{\beta}}_j, \max_{j \in \widehat{\tilde{\mathcal{L}}}_{rel}} \widehat{\tilde{\beta}}_j\right].$$

5.2 General k_z

The extension to the definition of the *FAS* for a general number of instruments that can violate the exogeneity or exclusion assumption is obtained by considering all possible just-identified model specifications. There are $s_{k_z} = k_z 2^{k_z-1}$ such specifications, resulting in e.g. $s_2 = 4$, $s_3 = 12$ and $s_4 = 32$. This is established as follows. For an instrument Z_ℓ, $\ell \in \{1, \ldots, k_z\}$, let the set $A_{-\ell} = \{r : r \in \{1, \ldots, k_z\}, r \neq \ell\}$ and let $\mathcal{C}_{-\ell} = \{C_{-\ell,m}\}_{m=1}^{M}$ denote the collection of all subsets of $A_{-\ell}$. The number of elements in $C_{-\ell,m}$ range from $0, 1, \ldots, k_z - 1$, so $\mathcal{C}_{-\ell}$ includes $\emptyset$ and $A_{-\ell}$. There are $M = 2^{k_z-1}$ such subsets. Define as in (20) the linear projection error

$$Z_{\ell|C_{-\ell,m}} := Z_\ell - \boldsymbol{Z}_{C_{-\ell,m}}^T \boldsymbol{\phi}_{C_{-\ell,m}\ell},$$

then $Z_{\ell|C_{-\ell,m}}$ is the transformed just-identifying instrument in the model specification where the instruments $\boldsymbol{Z}_{C_{-\ell,m}}$ are included as controls and the instruments $\boldsymbol{Z}_{A_{-\ell}\setminus C_{-\ell,m}}$ are omitted from the instrument set. For each $\ell \in \{1, \ldots, k_z\}$ there are $M = 2^{k_z-1}$ such transformed just-identifying instruments and so the total number of just-identifying specifications is given by $s_{k_z} = k_z 2^{k_z-1}$.

As an example, for $k_z = 3$, $\ell = 1$ and $C_{-1,m} = \{3\}$, $Z_{1|3}$ is the just-identifying instrument Z_1 for the model with Z_3 included as control and Z_2 excluded from the instrument set. The full sequence for Z_1 is given by the $2^2 = 4$ just identifying instruments $\{Z_1, Z_{1|2}, Z_{1|3}, Z_{1|2,3}\}$. This applies to all 3 instruments, so there are a total of 12 just-identifying model specifications. For model specification (2) and (3), let $\boldsymbol{\gamma} = (0, 0, \gamma_3)^T$ and $\boldsymbol{\alpha} = (0, \alpha_2, 0)^T$ with $\gamma_3 \neq 0$ and $\alpha_2 \neq 0$. Here Z_1 is a valid instrument, but with correlated instruments, Z_3 needs to be included as a control for it to identify β. Therefore $Z_{1|3}$ is here the valid instrument that identifies β if it is relevant. If the latter is the case, then the *FAS* as constructed below will contain β.

Let $\tilde{\boldsymbol{Z}}$ be the s_{k_z}-vector of transformed just-identifying instruments

$$\begin{aligned} \tilde{\boldsymbol{Z}} &= \left(Z_{\ell|C_{-\ell,m}}\right)_{\ell=1,\ldots,k_z; m=1,\ldots,2^{k_z-1}} \\ &= \left(\tilde{Z}_j\right)_{j=1,\ldots,s_{k_z}}. \end{aligned} \tag{23}$$

Then define as in (22) the s_{k_z}-vectors $\tilde{\boldsymbol{\pi}}$ and $\tilde{\boldsymbol{\psi}}$ with elements

$$\tilde{\pi}_j = \left(\mathrm{var}\left(\tilde{Z}_j\right)\right)^{-1} \mathrm{cov}\left(\tilde{Z}_j, X\right); \; \tilde{\psi}_j = \left(\mathrm{var}\left(\tilde{Z}_j\right)\right)^{-1} \mathrm{cov}\left(\tilde{Z}_j, Y\right), \tag{24}$$

for $j = 1, \ldots, s_{k_z}$.

It follows from model specification (2) and (3) that

$$\tilde{\psi}_j - \tilde{\pi}_j \beta = \left(\mathrm{var}\left(\tilde{Z}_j\right)\right)^{-1} \mathrm{cov}\left(\tilde{Z}_j, \tilde{U}\right),$$

where $\tilde{U} = Y - X\beta = \boldsymbol{Z}^T \boldsymbol{\gamma} + U$. We can now make the general partial exogeneity and exclusion assumption,

Assumption 9. *General partial exogeneity and exclusion: There are known constants $\tilde{\delta}_j \geq 0$ such that $\left|\left(var\left(\tilde{Z}_j\right)\right)^{-1} cov\left(\tilde{Z}_j, \tilde{U}\right)\right| \leq \tilde{\delta}_j$ for all $j \in \{1, \ldots, s_{k_z}\}$.*

It then follows from the arguments given in Sections 3, 4 and 5.1 that under Assumptions 1, 2, 5 and 9, the identified set for β is given by

$$\mathcal{B}\left(\tilde{\boldsymbol{\delta}}\right) = \left\{b \in \mathbb{R} : -\tilde{\boldsymbol{\delta}} \leq \left(\tilde{\boldsymbol{\psi}} - \tilde{\boldsymbol{\pi}} b\right) \leq \tilde{\boldsymbol{\delta}}\right\}.$$

The set of relevant instruments is specified as

$$\tilde{\mathcal{L}}_{rel} = \{j \in \{1, \ldots, s_{k_z}\} : \tilde{\pi}_j \neq 0\}, \tag{25}$$

and the falsification frontier is given by

$$FF = \left\{\tilde{\boldsymbol{\delta}}(b) \in \mathbb{R}^{s_{k_z}}_{\geq 0} : \tilde{\delta}_j(b) = \left|\tilde{\psi}_j - b\tilde{\pi}_j\right|, \, j = 1, \ldots, s_{k_z}, \, b \in \left[\min_{j \in \tilde{\mathcal{L}}_{rel}} \frac{\tilde{\psi}_j}{\tilde{\pi}_j}, \max_{j \in \tilde{\mathcal{L}}_{rel}} \frac{\tilde{\psi}_j}{\tilde{\pi}_j}\right]\right\}.$$

Then the generalized falsification adaptive set is

$$\begin{aligned} FAS &= \left[\min_{j \in \tilde{\mathcal{L}}_{rel}} \frac{\tilde{\psi}_j}{\tilde{\pi}_j}, \max_{j \in \tilde{\mathcal{L}}_{rel}} \frac{\tilde{\psi}_j}{\tilde{\pi}_j}\right] \\ &= \left[\min_{j \in \tilde{\mathcal{L}}_{rel}} \tilde{\beta}_j, \max_{j \in \tilde{\mathcal{L}}_{rel}} \tilde{\beta}_j\right], \end{aligned}$$

where, for $j \in \tilde{\mathcal{L}}_{rel}$, the IV estimands are given by

$$\tilde{\beta}_j = \frac{\tilde{\psi}_j}{\tilde{\pi}_j} = \frac{\mathrm{cov}\left(\tilde{Z}_j, Y\right)}{\mathrm{cov}\left(\tilde{Z}_j, X\right)} = \beta + \frac{\mathrm{cov}\left(\tilde{Z}_j, \tilde{U}\right)}{\mathrm{cov}\left(\tilde{Z}_j, X\right)}.$$

We summarize our main result, the generalization of the falsification adaptive set, in the following proposition. This generalizes Theorem 2 in MP, with the further result that this *FAS* satisfies Condition 1. This latter follows straightforwardly, as for each instrument all possible permutations of invalidity of the other instruments under Assumption 5 are considered. Hence, if there is at least one valid instrument, then at least one of these permutations will identify β, as long as the associated transformed just-identifying instrument is relevant.

Proposition 1. *For the linear model specification (2) and (3), suppose Assumptions 1, 2, 5 and 9 hold and the distribution of $(Y, X, \boldsymbol{Z})$ is known. i) The falsification adaptive set is given by*

$$FAS = \left[\min_{j \in \tilde{\mathcal{L}}_{rel}} \frac{\tilde{\psi}_j}{\tilde{\pi}_j}, \max_{j \in \tilde{\mathcal{L}}_{rel}} \frac{\tilde{\psi}_j}{\tilde{\pi}_j} \right], \tag{26}$$

where $\tilde{\psi}_j$ and $\tilde{\pi}_j$ are defined in (24) for $j = 1, \ldots, s_{k_z}$, with $s_{k_z} = k_z 2^{k_z - 1}$ the number of all possible just-identified IV model specifications. The $\tilde{Z}_j$ in (24) are the s_{k_z} just-identifying transformed instruments as defined in (23). An instrument $\tilde{Z}_j$ is relevant if $\tilde{\pi}_j \neq 0$, and $\tilde{\mathcal{L}}_{rel}$ is the set of relevant instruments as defined in (25). ii) A sufficient condition for the FAS to contain β and thus be the identified set for β is that there is at least one instrument that is relevant and valid.

For estimation, the same procedure is followed as in Section 5.1.2. The set of relevant instruments is estimated by

$$\widehat{\tilde{\mathcal{L}}}_{rel} = \left\{ j \in \{1, \ldots, s_{k_z}\} : \tilde{F}_j \geq C_n \right\},$$

where $\tilde{F}_j$ is the F-statistic for testing $H_0 : \tilde{\pi}_j = 0$ in the first-stage linear specification $\boldsymbol{x} = \tilde{\boldsymbol{z}}_j \tilde{\pi}_j + \boldsymbol{v}_j$, where $\tilde{\boldsymbol{z}}_j$ is the j-th column of the $n \times s_{k_z}$ matrix of just-identifying transformed instruments

$$\begin{aligned} \tilde{\boldsymbol{Z}} &= \left[\boldsymbol{M}_{Z_{C_{-\ell,m}}} \boldsymbol{z}_\ell \right]_{\ell=1,\ldots,k_z, m=1,\ldots,2^{k_z-1}} \\ &= [\tilde{\boldsymbol{z}}_j] \, . \end{aligned}$$

Let $\widehat{\tilde{\beta}}_j$ be the IV estimators

$$\widehat{\tilde{\beta}}_j = \frac{\tilde{\boldsymbol{z}}_j^T \boldsymbol{y}}{\tilde{\boldsymbol{z}}_j^T \boldsymbol{x}} = \frac{\widehat{\tilde{\psi}}_j}{\widehat{\tilde{\pi}}_j},$$

then the consistent estimator of FAS is given by

$$\widehat{FAS} = \left[\min_{j \in \widehat{\tilde{\mathcal{L}}}_{rel}} \widehat{\tilde{\beta}}_j, \max_{j \in \widehat{\tilde{\mathcal{L}}}_{rel}} \widehat{\tilde{\beta}}_j \right].$$

6 Some Empirical Examples

6.1 An Empirical Example with Two Instruments

MP20, Section 4.2, calculate $\widehat{FAS}_{excl}$ for the cross-country study by Alesina et al. (2013) on the origin of gender roles. We will focus here on the results for the model with the outcome variable "Female labor force participation in 2000". Their endogenous treatment variable is the "estimated proportion of citizens with ancestors that traditionally used the plough in pre-industrial agriculture" and they use geo-climatic area conditions for suitability for growing "plough-positive" crops and "plough-negative" crops as two instruments. The full estimation results are given in Table VIII in Alesina et al. (2013), which also gives details of other control variables that have been included in the model. In MP20, these results and the $\widehat{FAS}_{excl}$ are reported in their Table 4.

Table 1 reports the results for the model specification without continent fixed effects. The 2sls estimation results using both instruments show a significant negative effect of traditional plough use on female labor force participation, but this model is falsified with the p-value of the Hansen (1982) J-test being equal to 0.0009. With Z_1 being the "plough-positive environment" instrument, and Z_2 the "plough-negative environment" one, it is found that the F-statistic for $H_0 : \pi_2 = 0$ is equal to 0.95 and hence $Z_{2|1}$ is not in the relevant set. As reported in MP20, the F-statistic for $H_0 : \pi_1 = 0$ is equal to 78.20 and so it follows that the $\widehat{FAS}_{excl} = -14.31$, which is the IV estimate using $Z_{1|2}$ as the just-identifying and relevant instrument. So, although the two instruments produce significantly different results, falsifying the model, the $\widehat{FAS}_{excl}$ is a single point. As mentioned before, the falsification adaptive set should expand to account for the uncertainty about which assumption along the frontier is true, which is not the case here. For the FAS_{exo}, we get $\widehat{FAS}_{exo} = [-46.98, -22.65]$, which is further expanded to the estimate of FAS, $\widehat{FAS} = [-46.98, -14.31]$, which is the range of IV estimates of the three just-identifying relevant instruments, $Z_{1|2}$, Z_1 and Z_2. $\widehat{FAS}$ fully shows the uncertainty of the results due to the falsification of the model.

Table 1 also provides the estimates of $\tilde{\psi}_j$ and $\tilde{\pi}_j$, for $j = 1, \ldots, 4$. It shows that

Table 1: Country level IV estimation results, Alesina et al. (2013, Table VIII, column 1)

	Instruments				
	Z_1, Z_2	$Z_{1\|2}$	$Z_{2\|1}$	Z_1	Z_2
Traditional plough use	-21.63	-14.31	159.6	-22.65	-46.98
	(5.25)	(5.15)	(176.9)	(5.32)	(10.77)
F	40.21	78.20	0.95	74.01	19.00
J-test p-value	0.0009				
$\widehat{FAS}_{excl}$	-14.31				
$\widehat{FAS}_{exo}$	$[-46.98, -22.65]$				
$\widehat{FAS}$	$[-46.98, -14.31]$				
$\tilde{\psi}$		-10.64	18.93	-16.08	29.66
		(3.82)	(6.51)	(3.50)	(6.28)
$\tilde{\pi}$		0.744	0.119	0.710	-0.631
		(0.084)	(0.122)	(0.083)	(0.144)

Notes: Outcome variable "Female labor force participation rate in 2000", $n = 160$, 2sls estimation results for model without continent fixed effects. Heteroskedasticity robust test statistics and (standard errors). Z_1 is instrument "Plough-pos. environment", Z_2 is "Plough-neg. environment", $Z_{s|r}$ is just identifying instrument Z_s when Z_r is added as a control.

$\hat{\tilde{\pi}}_2 \,(= \widehat{\pi}_2)$ and $\hat{\tilde{\pi}}_4 \,(= \widehat{\pi}_2^*)$ have opposite signs. Therefore, with $\widehat{\beta}_{2sls} = \sum_{\ell=1}^{2} w_\ell \widehat{\beta}_\ell = \sum_{\ell=1}^{2} w_\ell \widehat{\beta}_\ell^*$, as shown in Sections 3.3 and 4.2, it follows that the weight w_2 is negative. The weights are here given by $w_1 = 1.042$ and $w_2 = 1 - w_1 = -0.042$. Note that using any pair of the four just-identifying instruments results in the same 2sls estimator, which is then of course also a linear combination of the associated pair of the just-identified estimators, with the weights adding up to 1. For example, when using $Z_{1|2}$ and Z_1 as the instruments, we have here that $\widehat{\beta}_{2sls} = 0.12\widehat{\beta}_1 + 0.88\widehat{\beta}_1^*$. For all the six pairs, we find that[1]

$$\begin{aligned} \widehat{\beta}_{2sls} &= 1.042\widehat{\beta}_1 - 0.042\widehat{\beta}_2 = 1.042\widehat{\beta}_1^* - 0.042\widehat{\beta}_2^* \\ &= 0.12\widehat{\beta}_1 + 0.88\widehat{\beta}_1^* = 0.12\widehat{\beta}_2 + 0.88\widehat{\beta}_2^* \\ &= 0.78\widehat{\beta}_1 + 0.22\widehat{\beta}_2^* \\ &= 0.006\widehat{\beta}_2 + 0.994\widehat{\beta}_1^*. \end{aligned}$$

Use of any of any of these pairs of instruments also results in the same value for the J-statistic, which is equivalent to a minimum-distance criterion, see Windmeijer (2019). The small p-value therefore indicates that the estimands, for example β_1^* and β_2^*, are different from each other.

Although $Z_{2|1}$ is found to be an irrelevant instrument, this is not the case for Z_2

[1]A proof for the equivalence of the weights for $\widehat{\beta}_1, \widehat{\beta}_1^*$ and $\widehat{\beta}_2, \widehat{\beta}_2^*$ is provided in the Appendix.

when used as the just-identifying instrument by itself. This is due to the strong negative correlation between the two instruments, with the (partial) sample correlation between them given by $\widehat{\rho}_{12} = -0.54$. Therefore if Z_2 is a valid instrument, with $\gamma_2 = \alpha_2 = 0$, it can identify β if Z_1 is invalid due to a violation of the exogeneity assumption, $\gamma_1 = 0$, $\alpha_1 \neq 0$, even when $\pi_2 = 0$, as long as $\rho_{12} \neq 0$. But clearly, when $\pi_2 = 0$, Z_2 cannot identify β if Z_1 instead violates the exclusion assumption.

6.2 An Empirical Example with Three Instruments

One of the main examples in MP is the empirical analysis of roads and trade by Duranton et al. (2014). The outcome variable is a measure of how much a city exports. The one considered in MP is called the "propensity to export weight". The treatment variable is the log number of kilometers of interstate highway within a city in 2007. Duranton et al. (2014) estimate the causal effect of within city highways on the propensity to export weight using instrumental variables. There are three potential instruments: $Z_1 = Plan$ is the log number of kilometers of highway in the city according to a planned highway construction map, approved by the federal government in 1947; $Z_2 = Railroads$ is the log number of kilometers of railroads in the city in 1898; and $Z_3 = Exploration$ is a measure of the quantity of historical exploration routes that passed through the city. For a fuller description see Duranton et al. (2014) and MP.

We will focus here on the model specification and estimation results as displayed in column 2 of Table 5 in Duranton et al. (2014) and in column 2 of Table I in MP. This model specification includes the additional control variables "log employment" and "Market access (export)". Estimation results for the two additional control variables have been omitted and notation-wise these variables have been partialled out. The first column in Table 2 replicates the 2sls estimation results using all 3 instruments. This specification is falsified by the J-statistic, which has a p-value of 0.043. The next columns give the estimation results for all twelve just-identified model specifications. The instruments $Z_{2|1,3}$ and $Z_{2|1}$ are found to be not relevant, and the resulting estimates of the falsification adaptive sets are given by $\widehat{FAS}_{excl} = [-0.32, 0.28]$ as in MP, $\widehat{FAS}_{exo} = [0.13, 1.09]$ and $\widehat{FAS} = [-0.61, 1.18]$.

The $\widehat{FAS}$ has quite a wide range of values here, as it correctly takes into account possible violations of exogeneity and exclusion restrictions. If for example $\boldsymbol{\gamma} = (0, 0, \gamma_3)^T$ and $\boldsymbol{\alpha} = (\alpha_1, 0, 0)^T$ with $\gamma_3 \neq 0$ and $\alpha_1 \neq 0$, then $Z_{2|3}$ is a valid instrument. It is found to be a relevant instrument, and the associated IV estimate is given by 1.18, which is the largest coefficient estimate of all just-identified specifications. $Z_{2|1,3}$ and $Z_{2|1}$ are found to be not relevant, with their F-statistics less than 10. In contrast, Z_2 as well as $Z_{2|3}$ are found to be relevant. As in the empirical example of Section 6.1, this can be explained by the strong correlation of the instruments. The sample partial correlation coefficients are here given by $\widehat{\rho}_{12} = 0.57$, $\widehat{\rho}_{13} = 0.34$ and $\widehat{\rho}_{23} = 0.11$.

The 2sls estimator is again the same when using any combination of three instruments from the twelve transformed just-identifying ones, as long as all indices $\{1, 2, 3\}$ are involved, as in for example $\{Z_1, Z_{1|2}, Z_{1|3}\}$. The 2sls weights for $\{Z_1, Z_2, Z_3\}$ and $\{Z_{1|2,3}, Z_{2|1,3}, Z_{3|1,2}\}$ are here given by $w = \{0.757, 0.126, 0.117\}$.

Table 2: IV estimation results, Duranton et al. (2014, Table 5, column 2)

	Instruments						
	Z_1, Z_2, Z_3	$Z_{1\|2,3}$	$Z_{2\|1,3}$	$Z_{3\|1,2}$	Z_1	Z_2	Z_3
log highway km	0.57	0.28	3.16	−0.32	0.55	1.09	0.13
	(0.16)	(0.25)	(1.39)	(0.86)	(0.17)	(0.26)	(0.38)
F	90.30	58.13	6.97	20.00	154.5	35.84	15.97
J-test p-value	0.043						
		$Z_{1\|2}$	$Z_{1\|3}$	$Z_{2\|1}$	$Z_{2\|3}$	$Z_{3\|1}$	$Z_{3\|2}$
log highway km		0.22	0.40	3.74	1.18	−0.61	−0.02
		(0.21)	(0.16)	(1.90)	(0.26)	(1.11)	(0.38)
F		81.14	122.45	5.29	34.31	14.27	31.07
$\widehat{FAS}_{excl}$	$[-0.32, 0.28]$						
$\widehat{FAS}_{exo}$	$[0.13, 1.09]$						
$\widehat{FAS}$	$[-0.61, 1.18]$						

Notes: Outcome variable "propensity to export weight", $n = 66$. Additional controls "log employment" and "Market access (export)". Heteroskedasticity robust test statistics and (standard errors). Z_1 is instrument "Plan", Z_2 is "Railroads", Z_3 is "Exploration".

6.2.1 Instrument Selection

When a model is falsified, the $\widehat{FAS}$ gives the range of estimates of all just-identified model specifications where the just-identifying instruments are found to be relevant. If a relevant instrument is valid, then the population *FAS* will contain β under our stated assumptions. As such, estimation of the *FAS* is not a selection tool. In the exclusion assumption violation only setting, as $Z_{2|1,3}$ is found to be an irrelevant instrument, MP however argued to automatically include $Z_2 = $ *Railroads*, as a control. If Z_1 and Z_3 are valid instruments and Z_2 violates the exclusion assumption, then the resulting 2sls estimator would of course be a consistent estimator of β, and the null hypothesis for the J-test would be true. If Z_2 violates the exogeneity assumption instead, it should clearly be omitted from the instrument set. However, instrument selection like that requires the assumption in this case that two instruments are valid. Then the valid instruments can be found by selecting the specification for which the J-test does not reject. The MP approach for hard thresholding with exclusion violation only is then the same as the approach of Guo et al. (2018) and Windmeijer et al. (2021).

Table 3: Results for Estimation with Two Instruments

	Instruments					
	Z_1, Z_2	$Z_{1\|3}, Z_{2\|3}$	Z_1, Z_3	$Z_{1\|2}, Z_{3\|2}$	Z_2, Z_3	$Z_{2\|1}, Z_{3\|1}$
log highway km	0.61	0.73	0.51	0.17	0.74	0.84
	(0.17)	(0.20)	(0.16)	(0.19)	(0.19)	(0.64)
F	91.64	81.13	102.27	65.34	38.14	11.87
J-test p-value	0.014	0.020	0.25	0.51	0.039	0.019

Notes: See notes to Table 2.

When considering the model specifications with two instruments only it seems natural to consider all possible combinations allowing for the violation of the exclusion or exogeneity assumption. The estimation results are presented in Table 3. Although $Z_{2|1,3}$ and $Z_{2|1}$ are not relevant instruments, we include them here as part of the pairs $\{Z_1, Z_2\}$, $\left\{Z_{1|3}, Z_{2|3}\right\}$ and $\left\{Z_{2|1}, Z_{3|1}\right\}$ as there is enough information in the data to reject these specifications. We find here that the J-statistics for the pairs $\{Z_1, Z_3\}$ and $\left\{Z_{1|2}, Z_{3|2}\right\}$ do not falsify the model specifications, with the p-values respectively given by 0.25 and 0.51. Hence there is no evidence against including *Railroads* as a control or dropping it from the instrument set. The estimation results are quite different in the two specifications, with an estimate of 0.51 (se 0.16) and $F = 102.3$ when Z_2 is dropped, and 0.17 (se 0.19) and $F = 65.34$ when Z_2 is included as a control. Selecting the model with the larger value of the F-statistic would then favor the model with Z_2 excluded. When adopting the selection strategy of choosing the model with the smallest value of the J-statistic, as in Andrews (1999), then the model with Z_2 included as a control would be selected, but there is no reason to specify this model as the automatic default for a falsified model with irrelevant instruments. And note that it could still be the case that Z_2 or $Z_{2|3}$ is the only valid and relevant instrument, as the assumption that in this case the majority of instruments is valid is not verifiable and a non-rejection of the J-test does not imply that instruments are valid.

To illustrate this further, Tables 4 and 5 present the same estimation results for the specification in Duranton et al. (2014) as presented in their Table 5, column 4, and in MP Table I, column 4 and Table II. Although this specification is not falsified by the J-test, MP do provide the estimate of FAS_{excl}. The results for the different falsification adaptive sets are here given by $\widehat{FAS}_{excl} = [0.18, 0.42]$, $\widehat{FAS}_{exo} = [0.34, 0.64]$ and $\widehat{FAS} = [0.18, 0.67]$. *Railroads* is again found to be an irrelevant instrument in the sense that the F-statistic for $Z_{2|1,3}$ is given by 1.27 and for $Z_{2|1}$ is given by 0.16. But Z_2 by itself and $Z_{2|3}$ are found to be relevant instruments due to the correlation structure. The sample partial correlation coefficients are given by $\widehat{\rho}_{12} = 0.63$, $\widehat{\rho}_{13} = 0.38$ and $\widehat{\rho}_{23} = 0.12$.

Table 4: IV estimation results, Duranton et al. (2014, Table 5, column 4)

	Instruments						
	Z_1, Z_2, Z_3	$Z_{1\|2,3}$	$Z_{2\|1,3}$	$Z_{3\|1,2}$	Z_1	Z_2	Z_3
log highway km	0.39	0.18	3.66	0.42	0.38	0.64	0.34
	(0.12)	(0.21)	(4.16)	(0.52)	(0.13)	(0.22)	(0.19)
F	84.71	54.75	1.27	26.95	129.7	40.75	23.79
J-test p-value	0.30						
		$Z_{1\|2}$	$Z_{1\|3}$	$Z_{2\|1}$	$Z_{2\|3}$	$Z_{3\|1}$	$Z_{3\|2}$
log highway km		0.21	0.40	8.78	0.67	0.26	0.34
		(0.15)	(0.16)	(22.3)	(0.24)	(0.54)	(0.19)
F		81.92	98.57	0.16	44.07	24.66	39.41
$\widehat{FAS}_{excl}$	$[0.18, 0.42]$						
$\widehat{FAS}_{exo}$	$[0.34, 0.64]$						
$\widehat{FAS}$	$[0.18, 0.67]$						

Notes: Outcome variable "propensity to export weight", $n = 66$. Additional controls "log employment", "Market access (export)", "log 1920 population", "log 1950 pop.", "log 2000 pop." and "log % manuf. emp.". Heteroskedasticity robust test statistics and (standard errors). Z_1 is instrument "Plan", Z_2 is "Railroads", Z_3 is "Exploration".

MP then again include *Railroads* by default as a control variable and presents the 2sls estimation results for that particular specification. In Table 5 we compare model specifications with 2 instruments, but ignore here any combination which would have an irrelevant instrument and therefore consider only the pairs $\{Z_1, Z_3\}$, $\left\{Z_{1|2}, Z_{3|2}\right\}$ and $\{Z_2, Z_3\}$. As the original specification was not falsified we don't expect these specifications to be falsified, as confirmed by the p-values of the J-statistics, which are respectively 0.83, 0.72 and 0.25. The model with largest value of the F-statistic and smallest value of the J-statistic in this case is the model with Z_2 excluded from the instrument set. The resulting 2sls estimate of 0.38 (se 0.12) is very close to the one based on all three instruments, 0.39 (se 0.12). The 2sls weights for the three instruments $\{Z_1, Z_2, Z_3\}$ are given by $w = \{0.81, 0.05, 0.14\}$, those for the two instruments $\{Z_1, Z_3\}$ are given by $w = \{0.98, 0.02\}$.

Table 5: Results for Estimation with Two Instruments

	Instruments		
	Z_1, Z_3	$Z_{1\|2}, Z_{3\|2}$	Z_2, Z_3
log highway km	0.38	0.23	0.51
	(0.12)	(0.14)	(0.16)
F	112.28	82.15	46.65
J-test p-value	0.83	0.72	0.25

Notes: See notes to Table 4.

7 Conclusions

We have generalized the falsification adaptive set of Masten and Poirier (2021) for the classical linear model with an endogenous variable, estimated by the method of instrumental variables with multiple instruments. It is the set of all just-identifying estimands where the just-identifying instrument is relevant. It reflects the model uncertainty when the baseline model is falsified, taking into account possible violations of both the exogeneity and exclusion assumptions and where an invalid instrument can violate either the exclusion or exogeneity assumption. It contains β if there is at least one valid and relevant instrument, and we recommend researchers to report estimates of this set when their baseline model is falsified.

Appendix

We show here that for the $k_z = 2$ case we have that $\widehat{\beta}_{2sls} = \tilde{\boldsymbol{w}}^T \left(\widehat{\beta}_1, \widehat{\beta}_1^*\right)^T = \tilde{\boldsymbol{w}}^T \left(\widehat{\beta}_2, \widehat{\beta}_2^*\right)^T$, as observed in Section 6.1.

For general instruments $\boldsymbol{Z} = [\boldsymbol{z}_1 \ \boldsymbol{z}_2]$ the 2sls weights on the instrument specific estimators $\widehat{\beta}_1^*$ and $\widehat{\beta}_2^*$ was found in Section 4.2 to be given by

$$w_\ell = \widehat{\pi}_\ell \boldsymbol{z}_\ell^T \boldsymbol{x} \left(\boldsymbol{x}^T \boldsymbol{P}_Z \boldsymbol{x}\right)^{-1},$$

for $\ell = 1, 2$, with $w_1 + w_2 = 1$, and where $\widehat{\pi}_\ell$ is the ℓ-th element of $\widehat{\boldsymbol{\pi}} = \left(\boldsymbol{Z}^T \boldsymbol{Z}\right)^{-1} \boldsymbol{Z}^T \boldsymbol{x}$,

$$\widehat{\boldsymbol{\pi}} = \begin{pmatrix} \widehat{\pi}_1 \\ \widehat{\pi}_2 \end{pmatrix} = \begin{pmatrix} \left(\boldsymbol{z}_{1|2}^T \boldsymbol{z}_{1|2}\right)^{-1} \boldsymbol{z}_{1|2}^T \boldsymbol{x} \\ \left(\boldsymbol{z}_{2|1}^T \boldsymbol{z}_{2|1}\right)^{-1} \boldsymbol{z}_{2|1}^T \boldsymbol{x} \end{pmatrix},$$

where

$$\boldsymbol{z}_{1|2} = \boldsymbol{M}_{z_2}\boldsymbol{z}_1 = \boldsymbol{z}_1 - \boldsymbol{z}_2\left(\boldsymbol{z}_2^T\boldsymbol{z}_2\right)^{-1}\boldsymbol{z}_2^T\boldsymbol{z}_1 = \boldsymbol{z}_1 - \boldsymbol{z}_2\widehat{\phi}_{21}$$

$$\boldsymbol{z}_{2|1} = \boldsymbol{M}_{z_1}\boldsymbol{z}_2 = \boldsymbol{z}_2 - \boldsymbol{z}_1\left(\boldsymbol{z}_1^T\boldsymbol{z}_1\right)^{-1}\boldsymbol{z}_1^T\boldsymbol{z}_2 = \boldsymbol{z}_2 - \boldsymbol{z}_1\widehat{\phi}_{12},$$

with $\widehat{\phi}_{21} = \left(\boldsymbol{z}_2^T\boldsymbol{z}_2\right)^{-1}\boldsymbol{z}_2^T\boldsymbol{z}_1$ and $\widehat{\phi}_{12} = \left(\boldsymbol{z}_1^T\boldsymbol{z}_1\right)^{-1}\boldsymbol{z}_1^T\boldsymbol{z}_2$.

Consider the instruments $\tilde{\boldsymbol{Z}}_1 = \left[\boldsymbol{z}_{1|2}\ \boldsymbol{z}_1\right]$, with instruments specific estimators respectively $\widehat{\beta}_1$ and $\widehat{\beta}_1^*$. Denote the associated 2sls weights $\tilde{\boldsymbol{w}}_1$, with first element given by

$$\tilde{w}_{1,1} = \tilde{\pi}_{1,1}\boldsymbol{z}_{1|2}^T\boldsymbol{x}\left(\boldsymbol{x}^T\boldsymbol{P}_{\tilde{Z}_1}\boldsymbol{x}\right)^{-1} = \tilde{\pi}_{1,1}\boldsymbol{z}_{1|2}^T\boldsymbol{x}\left(\boldsymbol{x}^T\boldsymbol{P}_Z\boldsymbol{x}\right)^{-1}$$

with the second equality following from the fact that $\boldsymbol{z}_{1|2}$ is a linear combination of $\boldsymbol{z}_1$ and $\boldsymbol{z}_2$. $\tilde{\pi}_{1,1}$ is the first element of $\tilde{\boldsymbol{\pi}}_1 = \left(\tilde{\boldsymbol{Z}}_1^T\tilde{\boldsymbol{Z}}_1\right)^{-1}\tilde{\boldsymbol{Z}}_1^T\boldsymbol{x}$.

We have that

$$\tilde{\boldsymbol{z}}_{1|2} = \boldsymbol{M}_{z_1}\boldsymbol{M}_{z_2}\boldsymbol{z}_1 = -\boldsymbol{z}_{2|1}\widehat{\phi}_{21}$$

and so

$$\tilde{\pi}_{1,1} = -\frac{\boldsymbol{z}_{2|1}^T\boldsymbol{x}}{\boldsymbol{z}_{2|1}^T\boldsymbol{z}_{2|1}\widehat{\phi}_{21}} = -\frac{\boldsymbol{z}_{2|1}^T\boldsymbol{x}}{\boldsymbol{z}_1^T\boldsymbol{z}_2\left(1 - \widehat{\phi}_{12}\widehat{\phi}_{21}\right)},$$

as

$$\boldsymbol{z}_{2|1}^T\boldsymbol{z}_{2|1}\widehat{\phi}_{21} = \left(\boldsymbol{z}_2^T\boldsymbol{z}_2 - \boldsymbol{z}_2^T\boldsymbol{z}_1\widehat{\phi}_{12}\right)\widehat{\phi}_{21} = \boldsymbol{z}_1^T\boldsymbol{z}_2\left(1 - \widehat{\phi}_{12}\widehat{\phi}_{21}\right),$$

and so

$$\tilde{\pi}_{1,1}\boldsymbol{z}_{1|2}^T\boldsymbol{x} = -\frac{\boldsymbol{x}^T\boldsymbol{z}_{2|1}\boldsymbol{z}_{1|2}^T\boldsymbol{x}}{\boldsymbol{z}_1^T\boldsymbol{z}_2\left(1 - \widehat{\phi}_{12}\widehat{\phi}_{21}\right)}.$$

Next consider the instruments $\tilde{\boldsymbol{Z}}_2 = \left[\boldsymbol{z}_{2|1}\ \boldsymbol{z}_2\right]$, with instruments specific estimators respectively $\widehat{\beta}_2$ and $\widehat{\beta}_2^*$. Denote the associated 2sls weights $\tilde{\boldsymbol{w}}_2$, with first element given by

$$\tilde{w}_{2,1} = \tilde{\pi}_{2,1}\boldsymbol{z}_{2|1}^T\boldsymbol{x}\left(\boldsymbol{x}^T\boldsymbol{P}_{\tilde{Z}_2}\boldsymbol{x}\right)^{-1} = \tilde{\pi}_{2,1}\boldsymbol{z}_{2|1}^T\boldsymbol{x}\left(\boldsymbol{x}^T\boldsymbol{P}_Z\boldsymbol{x}\right)^{-1},$$

with $\tilde{\pi}_{2,1}$ the first element of $\tilde{\boldsymbol{\pi}}_2 = \left(\tilde{\boldsymbol{Z}}_2^T\tilde{\boldsymbol{Z}}_2\right)^{-1}\tilde{\boldsymbol{Z}}_2^T\boldsymbol{x}$. By the derivations as above we get,

$$\tilde{\pi}_{2,1}\boldsymbol{z}_{2|1}^T\boldsymbol{x} = -\frac{\boldsymbol{x}^T\boldsymbol{z}_{1|2}\boldsymbol{z}_{2|1}^T\boldsymbol{x}}{\boldsymbol{z}_1^T\boldsymbol{z}_2\left(1 - \widehat{\phi}_{12}\widehat{\phi}_{21}\right)},$$

and so $\tilde{w}_{1,1} = \tilde{w}_{2,1}$. As $\tilde{w}_{1,2} = 1 - \tilde{w}_{1,1}$ and $\tilde{w}_{2,2} = 1 - \tilde{w}_{2,1}$, it follows that $\tilde{\boldsymbol{w}}_1 = \tilde{\boldsymbol{w}}_2 = \tilde{\boldsymbol{w}}$.

References

Alesina, A., P. Giuliano, and N. Nunn (2013): "On the Origins of Gender Roles: Women and the Plough," *The Quarterly Journal of Economics*, 128, 469–530.

Andrews, D. W. K. (1999): "Consistent Moment Selection Procedures for Generalized Method of Moments Estimation," *Econometrica*, 67, 543–563.

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.

Guo, Z., H. Kang, T. T. Cai, and D. S. Small (2018): "Confidence Intervals for Causal Effects with Invalid Instruments by Using Two-Stage Hard Thresholding with Voting," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 80, 793–815.

Hansen, L. P. (1982): "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, 50, 1029.

Kang, H., A. Zhang, T. T. Cai, and D. S. Small (2016): "Instrumental Variables Estimation With Some Invalid Instruments and its Application to Mendelian Randomization," *Journal of the American Statistical Association*, 111, 132–144.

Masten, M. A. and A. Poirier (2020): "Salvaging Falsified Instrumental Variable Models," *arXiv*, 1812.11598.

——— (2021): "Salvaging Falsified Instrumental Variable Models," *Econometrica*, 89, 1449–1469.

Windmeijer, F. (2019): "Two-Stage Least Squares as Minimum Distance," *The Econometrics Journal*, 22, 1–9.

Windmeijer, F., H. Farbmacher, N. Davies, and G. Davey Smith (2019): "On the Use of the Lasso for Instrumental Variables Estimation with Some Invalid Instruments," *Journal of the American Statistical Association*, 114, 1339–1350.

Windmeijer, F., X. Liang, F. P. Hartwig, and J. Bowden (2021): "The Confidence Interval Method for Selecting Valid Instrumental Variables," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 83, 752–776.