

A Framework for Eliciting, Incorporating, and Disciplining Identification Beliefs in Linear Models ^{*†}

Francis J. DiTraglia^{‡1} and Camilo García-Jimeno²

¹Department of Economics, University of Oxford

²Federal Reserve Bank of Chicago & NBER

Final Version: November 1, 2019, First Version: August 29, 2015

Abstract

To estimate causal effects from observational data, an applied researcher must impose beliefs. The instrumental variables exclusion restriction, for example, represents the belief that the instrument has no direct effect on the outcome of interest. Yet beliefs about instrument validity do not exist in isolation. Applied researchers often discuss the likely direction of selection and the potential for measurement error in their papers but lack formal tools for incorporating this information into their analyses. Failing to use all relevant information not only leaves money on the table; it runs the risk of leading to a contradiction in which one holds mutually incompatible beliefs about the problem at hand. To address these issues, we first characterize the joint restrictions relating instrument invalidity, treatment endogeneity, and non-differential measurement error in a workhorse linear model, showing how beliefs over these three dimensions are mutually constrained by each other and the data. Using this information, we propose a Bayesian framework to help researchers elicit their beliefs, incorporate them into estimation, and ensure their mutual coherence. We conclude by illustrating our framework in a number of examples drawn from the empirical microeconomics literature.

Keywords: Partial identification, Beliefs, Instrumental variables, Measurement error, Bayesian econometrics

JEL Codes: C10, C11, C18, C26

*The views expressed in this article are those of the authors and do not necessarily reflect the position of the Federal Reserve Bank of Chicago or the Federal Reserve System.

[†]We thank two anonymous referees, Daron Acemoglu, Thorsten Drautzburg, Richard Hahn, Hidehiko Ichimura, Laura Liu, Ulrich Müller, Frank Schorfheide, and Ben Ukert, as well as seminar participants at Princeton, Penn State, the Philadelphia FRB, the 2015 NSF-NBER SBIES, the 2015 MEG Meetings, and the 2016 ISBA World Meeting for helpful comments and suggestions. We thank Mallick Hossain and Alejandro Sánchez for excellent research assistance and acknowledge support from a UPenn URF award.

[‡]Corresponding Author: francis.ditraglia@economics.ox.ac.uk, Manor Road, Oxford OX1 3UQ, UK.

“Belief is so important! A hundred contradictions might be true.”

— Blaise Pascal, *Pensées*

1 Introduction

To identify causal effects from observational data, an applied researcher must augment the data with her beliefs. The exclusion restriction in an instrumental variables (IV) regression, for example, represents the belief that the instrument has no direct effect on the outcome of interest. Even when this belief cannot be tested directly, applied researchers know how to think about it and how to debate it. In practice, however, not all beliefs are treated equally. In addition to “formal beliefs” such as the IV exclusion restriction – beliefs that are directly imposed to obtain identification – researchers often state a number of “informal beliefs.” While not directly imposed on the problem, informal beliefs play an important role in interpreting results and reconciling conflicting estimates. Papers that report IV estimates, for example, almost invariably state the authors’ belief about the sign of the correlation between the endogenous treatment and the error term but do not exploit this information in estimation.¹ Another common informal belief concerns the extent of measurement error. When researchers observe an ordinary least squares (OLS) estimate that is substantially smaller than, but has the same sign as its IV counterpart, classical measurement error, with its attendant “least squares attenuation bias,” is often suggested as the likely cause.

Relegating informal beliefs to second-class status is both wasteful of information and dangerous; beliefs along different dimensions of the problem are mutually constrained by each other, the model, and the data. By failing to explicitly incorporate all relevant information, applied researchers both leave money on the table and, more importantly, risk reasoning to a contradiction by expressing mutually incompatible beliefs. Although this point is general, we illustrate its implications here in the context of a linear model

$$y = \beta T^* + \mathbf{x}'\boldsymbol{\gamma} + u \tag{1}$$

$$T^* = \pi z + \mathbf{x}'\boldsymbol{\eta} + v \tag{2}$$

$$T = T^* + \tilde{w} \tag{3}$$

where T^* is a potentially endogenous treatment, y is an outcome of interest, and \mathbf{x} is a vector

¹Referring to more than 60 papers published in the top three empirical journals between 2002 and 2005, Moon and Schorfheide (2009) note that “in almost all of the papers the authors explicitly stated their beliefs about the sign of the correlation between the endogenous regressor and the error term; yet none of the authors exploited the resulting inequality moment condition in their estimation.”

of exogenous controls. Our goal is to estimate the causal effect of T^* on y , namely β , but we observe only T , a noisy measure of T^* polluted by measurement error \tilde{w} . While we are fortunate to have an instrument z at our disposal, it may not satisfy the exclusion restriction: z is potentially correlated with u . This scenario is typical in applied work: endogeneity is the rule rather than the exception, the treatments of greatest interest are often the hardest to measure, and the validity of a proposed instrument is almost always debatable.

We focus on two cases that are common in applied work. In the first T^* has no support restrictions and is subject to classical measurement error. In the second T^* is binary and thus any errors in measurement *must* be non-classical.² To accommodate both cases within a single framework, we derive our results under the assumption that \tilde{w} is *non-differential*. This permits correlation between \tilde{w} and T^* but imposes the restriction that \tilde{w} is uncorrelated with all other random variables in the system conditional on T^* . We begin by deriving the sharp identified set relating treatment endogeneity, instrument invalidity, and non-differential measurement error when T^* has unrestricted support. To the best of our knowledge, this result is new to the literature. Turning our attention to the binary T^* case, we then show that adding support restrictions provides additional identifying information via cross-parameter restrictions. In both cases, however, the data alone provide no restrictions on β . As such, the addition of researcher beliefs is unavoidable. Using our characterization of the identified set, we propose a framework for Bayesian inference for the treatment effect of interest that combines the data with researcher beliefs in a coherent and transparent way. As we show in our empirical examples, this framework not only allows researchers to incorporate relevant problem-specific beliefs, but helps them to refine and discipline them by revealing any inconsistencies that may be present.

Whenever one imposes information beyond what is contained in the data, it is crucial to make clear how this information affects the ultimate result. Accordingly, we decompose our problem into a vector of partially-identified structural parameters θ , and a vector of point-identified reduced form parameters φ . The vector θ contains the parameters that govern instrument invalidity, regressor endogeneity and measurement error, while φ contains observable moments obtained from reduced form regressions of (y, T, z) on \mathbf{x} . This decomposition is structured so that the data are only informative about θ through φ , revealing precisely how any identification beliefs we may choose to impose enter the problem.³ In particular, the data rule out certain values of φ , while our beliefs place restrictions on the conditional identified set $\Theta(\varphi)$ for θ . A prior over the conditional identified set $\Theta(\varphi)$ will

² If $T^* = 1$, the only way it can be mis-measured is downwards: $T = 0$. If $T^* = 0$ the only way it can be mis-measured is upwards: $T = 1$. Hence \tilde{w} must be negatively correlated with T^* .

³Such a decomposition is called a *transparent parameterization* in the statistics literature. See, for example [Gustafson \(2015\)](#).

never be updated by any amount of data. For this reason, prior elicitation for θ is particularly crucial. Our approach to elicitation for θ has two components. First, we parameterize measurement error, regressor endogeneity, and instrument invalidity in terms of intuitive, empirically meaningful parameters: correlations and what is in essence a signal-to-noise ratio. Second, because it can be challenging for researchers to articulate fully informative prior information, we consider only relatively weak prior beliefs in the form of sign and interval restrictions on the components of θ . These are fairly easy to elicit in practice and can be surprisingly informative about the causal effect of interest. We present two complementary approaches to Bayesian inference for the structural parameters: inference for the identified set Θ , and inference for the partially identified parameter θ under a conditionally uniform reference prior. We compare and contrast these approaches below.

While measurement error, treatment endogeneity, and invalid instruments have all generated voluminous literatures, to the best of our knowledge this is the first paper to carry out a partial identification exercise in which all three problems can be present simultaneously. Our main point is simple but has important implications for applied work that have been largely overlooked; measurement error, treatment endogeneity, and instrument invalidity are mutually constrained by each other and the data in a manner that can only be made apparent by characterizing the full identified set for the model. Because the dimension of this set is strictly smaller than the number of variables used to describe it, the constraints of the model could easily contradict prior researcher beliefs. Given the shape of the identified set, the belief that z is a valid instrument, for example, could imply an implausible amount of measurement error or a selection effect with the opposite of the expected sign. In this way our framework provides a means of reconciling and refining beliefs that would not be possible based on introspection alone. We are by no means the first to recognize the importance of requiring that beliefs be compatible. [Kahneman and Tversky \(1974\)](#), for example, make a closely related point in their discussion of heuristic decision-making under uncertainty. Even if specific probabilistic assessments appear coherent on their own,

an internally consistent set of subjective probabilities can be incompatible with other beliefs held by the individual ... For judged probabilities to be considered adequate, or rational, internal consistency is not enough. The judgements must be compatible with the entire web of beliefs held by the individual. Unfortunately, there can be no simple formal procedure for assessing the compatibility of a set of probability judgements with the judge's total system of beliefs (p. 1130).

Our purpose here is to take up the challenge laid down by [Kahneman and Tversky \(1974\)](#) and provide just such a formal procedure for assessing the compatibility of researcher beliefs over treatment endogeneity, measurement error, and instrument invalidity in linear models.

Although the intuition behind our procedure is straightforward, the details are more involved. For this reason we provide free and open-source software in R to make it easy for applied researchers to implement the methods described in this paper.⁴

This paper contributes to a small but growing literature on the Bayesian analysis of partially-identified models, including Poirier (1998), Richardson et al. (2011), Moon and Schorfheide (2012), Hahn et al. (2016), and Gustafson (2015). Some recent contributions to the literature on structural vector autoregression models (Amir-Ahmadi and Drautzburg, 2019; Arias et al., 2018; Baumeister and Hamilton, 2015) also explore related ideas. Because we discuss, as part of our exercise, Bayesian inferences for the identified set, our work relates to Kitagawa (2012), Kline and Tamer (2016), and Chen et al. (2016) who give sufficient conditions under which such inferences have a valid frequentist interpretation.

Our results relate to the classical literature on errors in variables in linear models, for example Klepper and Leamer (1984), Leamer (1987), and Bekker et al. (1987). The main distinction between our paper and this literature is threefold. First, our regressor of interest T^* is endogenous; second, the measurement error \tilde{w} that generates our observed regressor T may be non-classical; third we consider settings in which a (potentially imperfect) instrumental variable is available. While the proxy variable setting considered in Krasker and Pratt (1986) and Bollinger (2003) can be interpreted as a non-classical measurement error problem, these papers likewise consider only exogenous regressors and do not rely on an instrumental variable. Our results also relate to a large literature on estimating the effect of mis-measured binary regressors without relying on instrumental variables. An early contribution is Bollinger (1996) who provides partial identification bounds for the effect of an exogenous, binary regressor subject to non-differential mis-classification. van Hasselt and Bollinger (2012) derive additional bounds for the same model. Bollinger and van Hasselt (2017) propose a Bayesian inference procedure based on these bounds and consider an extension that addresses potential endogeneity in the true, unobserved regressor by placing a prior on its covariance with the error term. In contrast, Kreider and Pepper (2007), Kreider et al. (2012), and Gundersen et al. (2012) derive partial identification bounds for the effect of a binary regressor subject to *arbitrary* mis-classification error when the outcome of interest is also binary. The latter two papers allow for endogeneity in the true, unobserved regressor.

Because we consider a situation in which an instrumental variable is available, our setting is more closely related to that considered by Kane et al. (1999), Black et al. (2000), Frazis and Lowenstein (2003), Lewbel (2007), Mahajan (2006) and Hu (2008). The key lesson from these papers is that the two-stage least squares (TSLS) estimator is inconsistent even if the instrument is valid. When the treatment is exogenous, however, it is possible to construct a

⁴See <https://github.com/fditraglia/ivdoctr>.

non-linear method of moments estimator that recovers the treatment effect using a discrete instrumental variable. Unlike these papers, we consider a setting in which the binary treatment of interest may be endogenous. As shown in [DiTraglia and García-Jimeno \(2019\)](#) the usual instrumental variable assumption is insufficient to identify the effect of an endogenous, mis-measured, binary treatment. While that paper provides a point identification result under a stronger instrument exclusion restriction, we do not rely on it here. Instead we allow for an *invalid* instrument and derive partial identification bounds.

Two papers that similarly consider partial identification under instrument invalidity are [Conley et al. \(2012\)](#) and [Nevo and Rosen \(2012\)](#). Like us, [Conley et al. \(2012\)](#) adopt a Bayesian approach that allows for a violation of the IV exclusion restriction, but they do not explore the relationship between treatment endogeneity and instrument invalidity. In contrast, [Nevo and Rosen \(2012\)](#) derive bounds for a causal effect in the setting where an endogenous regressor is “more endogenous” than the variable used to instrument it is invalid. Our framework encompasses the settings considered in these two papers, but is strictly more general in that we allow for measurement error simultaneously with treatment endogeneity and instrument invalidity. More importantly, the central message of our paper is that it can be misleading to impose beliefs on only one dimension of a partially identified problem unless one has a way of ensuring their mutual consistency with all other relevant researcher beliefs. For example, although a single valid instrument solves both the problem of classical measurement error and treatment endogeneity, it is insufficient to carry out a partial identification exercise that merely relaxes the exclusion restriction, as in [Conley et al. \(2012\)](#). Values for the correlation between z and u that seem plausible when viewed in isolation could easily imply implausible amounts of measurement error or treatment endogeneity.

The remainder of this paper is organized as follows. [Section 2](#) derives the sharp identified set when T^* has unrestricted support. [Section 3](#) considers the case in which T^* is binary, deriving additional cross-parameter restrictions that apply in this setting. [Section 4](#) details our two approaches to Bayesian inference, including details of prior elicitation, using the results of [Sections 3](#) and [4](#). [Section 5](#) presents a number of substantive empirical examples illustrating our procedure in both the classical measurement error and binary T^* cases, and [Section 6](#) concludes. Proofs, auxiliary results, and additional computational details appear in an online appendix.

2 The Identified Set

In this section we derive the joint restrictions relating measurement error, regressor endogeneity, and instrument invalidity given the observed data. We then use these restrictions

to show how the identified set for β depends on researcher beliefs over the three dimensions. Our approach is as follows. First, we use the assumption of non-differential measurement error to re-write (3) in terms of a *classical* measurement error component w and a parameter ψ that governs the “non-classical” part of measurement error, an approach similar to that followed by [Bollinger \(2003\)](#) in a proxy-variable setting.

Second, we relate the structural model from (1)–(3) to a system of reduced form regressions of (y, T, z) on \mathbf{x} . The restrictions that we use in our partial identification exercise below arise from the mapping between structural and reduced form covariance matrices, along with the assumption of non-differential measurement error. Third, we re-parameterize our problem to “absorb” the non-classical measurement error parameter ψ . This allows us to proceed *as though* the measurement error were classical, and adjust for ψ in a second step, greatly simplifying the calculations. The bounds we derive in this section are sharp provided that T^* has full support. When the support of T^* is restricted, however, it may be possible to tighten them, a possibility that we explore for a binary T^* in [section 3](#) below.

2.1 Model and Assumptions

We begin by stating the basic assumptions that will be used throughout the paper.

Assumption 2.1 (Model). *We observe (y, T, z, \mathbf{x}) generated from (1)–(3), where*

- (i) \mathbf{x} is exogenous: $\text{Cov}(\mathbf{x}, u) = \mathbf{0}$;
- (ii) v is a projection error: $\text{Cov}(\mathbf{x}, v) = \mathbf{0}$ and $\text{Cov}(z, v) = 0$;
- (iii) z is relevant for T^* : $\pi \neq 0$;
- (iv) \mathbf{x} includes a constant, so that $\mathbb{E}[u] = \mathbb{E}[v] = 0$;
- (v) T is positively correlated with T^* : $\text{Cov}(T, T^*) > 0$.

The only substantive restrictions in [Assumption 2.1](#) are (i) and (v): (i) assumes that the control regressors \mathbf{x} are exogenous, while (v) assumes that the mis-measured regressor T is positively correlated with the true, unobserved regressor T^* . [Assumption 2.1](#) (ii) can be taken as the *definition* of the error term v from (2). It equals the residual from a projection of the unobserved regressor of interest T^* on the instrument z and exogenous control regressors \mathbf{x} . [Assumption 2.1](#) (iii) is the standard instrumental variables relevance condition, but stated for the unobserved true regressor T^* rather than the observed, mis-measured regressor T . Although T^* is unobserved, [Assumption 2.1](#) (iii) is testable under our other assumptions.⁵ Throughout this paper we will abstract from weak instrument considerations.

⁵See (9) and the discussion immediately following it for details.

The main additional assumption that we rely on below concerns the nature of the measurement error \tilde{w} from (3).

Assumption 2.2 (Non-differential Measurement Error).

$$\begin{bmatrix} \text{Cov}(u, \tilde{w}) \\ \text{Cov}(z, \tilde{w}) \\ \text{Cov}(\mathbf{x}, \tilde{w}) \end{bmatrix} = \psi \begin{bmatrix} \text{Cov}(u, T^*) \\ \text{Cov}(z, T^*) \\ \text{Cov}(\mathbf{x}, T^*) \end{bmatrix}, \quad \psi \equiv \frac{\text{Cov}(T^*, \tilde{w})}{\text{Var}(T^*)}.$$

Assumption 2.2 requires that any correlation between \tilde{w} and (u, z, \mathbf{x}) arises solely from correlation between T^* and (u, z, \mathbf{x}) . In other words we assume that T contains no *additional information* about (u, z, \mathbf{x}) beyond that contained in T^* . Non-differential measurement error is the natural generalization of classical measurement error to settings where T and T^* have restricted support. As such, it is widely used in the literature on mis-classified discrete variables (e.g. DiTraglia and García-Jimeno, 2019; Frazis and Lowenstein, 2003; Hu, 2008; Lewbel, 2007; Mahajan, 2006). When $\psi = 0$, **Assumption 2.2** reduces to the classical case. When $\psi \neq 0$ it generalizes classical measurement error by allowing \tilde{w} to be correlated with T^* . This extra generality is necessary if we wish to consider a binary T^* because \tilde{w} must be correlated with T^* in this case: if $T^* = 1$ then \tilde{w} must be 0 or -1 ; if $T^* = 0$ then \tilde{w} must be 0 or 1. **Assumption 2.2** places no restriction on the conditional distribution of T given T^* and hence no restriction on ψ ; it merely imposes that T is *exogenous* after projecting out T^* . This is indeed a restriction, but a strictly weaker one than classical measurement error.

Before proceeding, we require some additional notation. First let

$$\tau \equiv \mathbb{E}[\tilde{w}] - \psi \mathbb{E}[T^*], \quad w \equiv \tilde{w} - \tau - \psi T^* \tag{4}$$

where ψ is as defined in **Assumption 2.2**. Using (4), we can re-write (3) as

$$T = \tau + (1 + \psi)T^* + w \tag{5}$$

where $(1 + \psi) > 0$ by **Assumption 2.1** (v), to ensure that T is positively correlated with T^* . Both (3) and (5) are completely without loss of generality: (3) can be viewed as the definition of \tilde{w} and (5) as the corresponding definition of w . Because w is defined as the residual from a projection of \tilde{w} onto T^* and a constant, it has zero mean and is uncorrelated with T^* by construction, making (5) more convenient to work with than (3). In contrast, \tilde{w} may have a non-zero mean and be correlated with T^* . Although T and T^* are positively correlated by **Assumption 2.1** (v), note that the correlation between T^* and \tilde{w} may be positive or negative as $\psi \in (-1, +\infty)$.

At the heart of our partial identification exercise is the relationship between reduced form and structural covariance matrices. Define the *reduced form* model as

$$y = \mathbf{x}'\boldsymbol{\varphi}_y + \varepsilon, \quad T = \mathbf{x}'\boldsymbol{\varphi}_T + \xi, \quad z = \mathbf{x}'\boldsymbol{\varphi}_z + \zeta \quad (6)$$

where $(\varepsilon, \xi, \zeta)$ are projection errors with covariance matrix

$$\Sigma \equiv \text{Var} \begin{bmatrix} \varepsilon \\ \xi \\ \zeta \end{bmatrix} = \begin{bmatrix} s_{11} & s_{12} & s_{13} \\ & s_{22} & s_{23} \\ & & s_{33} \end{bmatrix}. \quad (7)$$

Under [Assumption 2.1](#) (y, T, z, \mathbf{x}) are observed, so $\Phi \equiv (\boldsymbol{\varphi}_y, \boldsymbol{\varphi}_T, \boldsymbol{\varphi}_z)$ and Σ are point identified. Throughout the paper, we will refer to Σ as the *reduced form covariance matrix*. To avoid trivial but uninteresting cases, we assume throughout that Σ is positive definite. Let Ω denote the covariance matrix of (u, v, ζ, w) . We will refer to Ω as the *structural covariance matrix*.⁶ Ω is unobserved because T^* is unobserved and potentially endogenous. We assume that Ω is “well-behaved” in the following sense.

Assumption 2.3.

- (i) *The covariance matrix Ω of (u, v, ζ, w) exists and is finite.*
- (ii) *The covariance matrix Ω_{11} of (u, v, ζ) is positive definite.*

[Assumption 2.3](#) does not require that Ω be positive definite. This allows for the possibility that there is no measurement error, in which case $\text{Var}(w) = 0$. Note that we treat w rather than \tilde{w} as the “structural” measurement error. The advantage of following this convention is that w , unlike \tilde{w} , satisfies all of the assumptions of classical measurement error, as shown in the following lemma.

Lemma 2.1. *Under Assumptions [2.1](#), [2.2](#), and [2.3](#) (i), we have $\text{Cov}(\mathbf{x}, w) = \mathbf{0}$ and*

$$\Omega = \begin{bmatrix} \Omega_{11} & \mathbf{0} \\ \mathbf{0}' & \sigma_w^2 \end{bmatrix}, \quad \Omega_{11} = \begin{bmatrix} \sigma_u^2 & \sigma_{uv} & \sigma_{u\zeta} \\ \sigma_{uv} & \sigma_v^2 & 0 \\ \sigma_{u\zeta} & 0 & \sigma_\zeta^2 \end{bmatrix}. \quad (8)$$

[Equation 8](#) allows for the possibility that z is an invalid instrument, $\sigma_{u\zeta} \neq 0$, and that T^* is endogenous, $\sigma_{uv} \neq 0$. The zeros in Ω arise from [Assumption 2.1](#) (ii), which ensures that v is uncorrelated with ζ , and [Assumption 2.2](#), which ensures that w has the properties

⁶Note that our convention treats ζ as both a structural and reduced form error.

of classical measurement error. We now turn our attention to the relationship between the reduced form covariance matrix Σ and the structural covariance matrix Ω . This relationship emerges as a corollary of the following lemma.

Lemma 2.2. *Under Assumptions 2.1–2.3,*

$$\begin{aligned}\varepsilon &= \beta(\pi\zeta + v) + u & \varphi_y &= \beta(\pi\varphi_z + \eta) + \gamma \\ \xi &= (1 + \psi)(\pi\zeta + v) + w & \varphi_T &= \tau\mathbf{e}_1 + (1 + \psi)(\pi\varphi_z + \eta)\end{aligned}$$

where $\mathbf{e}_1 = (1, 0, \dots, 0)'$ denotes the first standard basis vector.

Lemma 2.2 shows that the reduced form coefficients φ_T and φ_y are functions of the structural parameters (β, π, ψ) . While it may appear from this result that knowledge of $(\varphi_y, \varphi_T, \varphi_z)$ provides additional identifying information, this is not the case. Given values for the reduced form regression coefficients $(\varphi_y, \varphi_T, \varphi_z)$, we can construct values of the structural regression coefficients η and γ that are consistent with *any* desired values of the other structural parameters, namely

$$\eta = \frac{\varphi_T - \tau\mathbf{e}_1}{1 + \psi} - \pi\varphi_z, \quad \gamma = \frac{\beta\varphi_T}{1 + \psi}$$

where **Assumption 2.1** (v) justifies division by $(1 + \psi)$: if $\text{Cov}(T, T^*) > 0$ then $\psi > -1$ as seen from (5). More importantly, **Lemma 2.2** implies that Σ is related to Ω according to

$$\Sigma = \Gamma\Omega\Gamma', \quad \Gamma \equiv \begin{bmatrix} 1 & \beta & \beta\pi & 0 \\ 0 & (1 + \psi) & (1 + \psi)\pi & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

Expanding $\Sigma = \Gamma\Omega\Gamma'$, we obtain the following:

$$s_{23} = (1 + \psi)\pi s_{33} \tag{9}$$

$$s_{13} = \sigma_{u\zeta} + \beta\pi s_{33} \tag{10}$$

$$s_{22} = (1 + \psi)^2 (\sigma_v^2 + \pi^2 s_{33}) + \sigma_w^2 \tag{11}$$

$$s_{12} = (1 + \psi) [(\sigma_{uv} + \pi\sigma_{u\zeta}) + \beta (\sigma_v^2 + \pi^2 s_{33})] \tag{12}$$

$$s_{11} = \sigma_u^2 + 2\beta (\sigma_{uv} + \pi\sigma_{u\zeta}) + \beta^2 (\sigma_v^2 + \pi^2 s_{33}). \tag{13}$$

Equations (9)–(13) constitute the restrictions that we will use to carry out our partial identification exercise below. **Equation 9** reveals that **Assumption 2.1** (iii), instrument relevance, is testable: $(1 + \psi)\pi = (s_{23}/s_{33})$ and $(1 + \psi)$ cannot equal zero by **Assumption 2.1** (v). As

shown in the following lemma, however, Assumptions 2.1–2.3 and the relationship $\Sigma = \Gamma\Omega\Gamma'$ impose no restrictions on the parameter ψ other than $\psi > -1$.

Lemma 2.3. *Suppose that the vector $\theta \equiv (\pi, \beta, \psi, \sigma_u, \sigma_v, \sigma_w, \sigma_{uv}, \sigma_{u\zeta})$ of structural parameter values satisfies Assumptions 2.1–2.3 and Equations 9–13. Then, for any $\psi' > -1$, so does $\theta' \equiv (\pi', \beta', \psi', \sigma_u, \sigma'_v, \sigma_w, \sigma'_{uv}, \sigma_{u\zeta})$ where we define*

$$\pi' \equiv \left(\frac{1 + \psi}{1 + \psi'} \right) \pi, \quad \beta' \equiv \left(\frac{1 + \psi'}{1 + \psi} \right) \beta, \quad \sigma'_v \equiv \left(\frac{1 + \psi}{1 + \psi'} \right) \sigma_v, \quad \sigma'_{uv} \equiv \left(\frac{1 + \psi}{1 + \psi'} \right) \sigma_{uv}.$$

Lemma 2.3 shows that, without further restrictions, the reduced form covariance matrix contains no information about ψ . Indeed an even stronger result holds: unless T^* has support restrictions, a model with structural parameters θ is *observationally equivalent* to one with structural parameters θ' .⁷ Intuitively, because T^* is unobserved we are free to arbitrarily re-scale both sides of (2) – effectively “redefining” T^* – so long as we absorb this rescaling into the remaining parameters of the system. If T^* has a restricted support, however, such an arbitrary rescaling is no longer possible. For example, if T^* is binary, certain choices of scale can be ruled out by observing the distribution of T . In this case it is still true that Σ on its own contains no information about ψ , but the binary nature of T^* creates additional cross-parameter restrictions that can be used to bound ψ . Because binary treatments are common in applied work, we develop this special case in full detail in section 3. Analogous reasoning applies to the parameter τ from (5). Without support restrictions on T^* we can shift τ arbitrarily while fixing $\mathbb{E}[T]$, absorbing the difference into $\mathbb{E}[T^*]$ and the first-stage intercept.

2.2 A Convenient Parameterization

Before proceeding to derive the joint restrictions between measurement error, regressor endogeneity, and instrument invalidity, we first re-write equations 9–13 in a form that simplifies both our mathematical derivations and, ultimately, the elicitation of researcher beliefs. To begin, we define a reduced form regression for the *unobserved* regressor T^* . Using logic analogous to that of Lemma 2.2, we can write

$$T^* = \mathbf{x}'\boldsymbol{\varphi}_T^* + \xi^*, \quad \boldsymbol{\varphi}_T^* = \pi\boldsymbol{\varphi}_z + \boldsymbol{\eta}, \quad \xi^* = \pi\zeta + v. \quad (14)$$

Since ζ is uncorrelated with v by Assumption 2.1 (ii), it follows that

$$\sigma_{u\xi^*} \equiv \text{Cov}(u, \xi^*) = \sigma_{uv} + \pi\sigma_{u\zeta}. \quad (15)$$

⁷See the proof of Theorem 2.1 for details.

Equation (15) shows that endogeneity in T^* arises from two sources: invalidity of the instrument z , and correlation between the error terms u and v . By representing regressor endogeneity in terms of $\sigma_{u\xi^*}$, (15) allows us to eliminate σ_{uv} from (9)–(13). Next we define the parameter κ as

$$\kappa \equiv \frac{\text{Var}(\xi^*)}{\text{Var}(\xi)} = \frac{\text{Var}(\pi\zeta + v)}{s_{22}} = \frac{\pi^2 s_{33} + \sigma_v^2}{s_{22}} = \left(\frac{1}{1 + \psi} \right)^2 \left(\frac{s_{22} - \sigma_w^2}{s_{22}} \right) \quad (16)$$

where the last equality follows by solving (11) for $(\pi^2 s_{33} + \sigma_v^2)$. In the special case where \mathbf{x} includes only a constant, T^* is exogenous, and the measurement error is classical, κ measures the degree of attenuation bias present in the OLS estimator. More generally, κ measures the proportion of “signal” contained in the reduced form error ξ^* . If $\kappa = 1/2$, for example, this means that half of the variation in ξ is generated by ξ^* , and the remainder is “noise” arising from w . Unlike σ_w^2 , κ has bounded support: $\kappa \in (0, 1]$. When $\kappa = 1$, $\sigma_w^2 = 0$ so there is no measurement error; the limit as κ approaches zero corresponds to taking σ_w^2 to its maximum possible value: s_{22} . Finally, define

$$\tilde{\beta} \equiv \frac{\beta}{1 + \psi}, \quad \tilde{\pi} \equiv (1 + \psi)\pi, \quad \tilde{\sigma}_v^2 \equiv (1 + \psi)^2 \sigma_v^2, \quad \tilde{\sigma}_{u\xi^*} \equiv (1 + \psi)\sigma_{u\xi^*}, \quad \tilde{\kappa} \equiv (1 + \psi)^2 \kappa. \quad (17)$$

The parameters defined in (17) correspond to setting $\psi' = 0$ in Lemma 2.3, which “absorbs” the non-classical component of measurement error, ψ , into the definitions of the remaining parameters. Note that if the measurement error \tilde{w} is in fact classical, then $\psi = 0$ so that $\tilde{\beta} = \beta$, $\tilde{\pi} = \pi$, and so on. Using (15)–(17), we can re-write (9)–(13) as

$$s_{23} = \tilde{\pi} s_{33} \quad (18)$$

$$s_{13} = \sigma_{u\zeta} + \tilde{\beta} \tilde{\pi} s_{33} \quad (19)$$

$$s_{22} = \tilde{\kappa} s_{22} + \sigma_w^2 \quad (20)$$

$$s_{12} = \tilde{\sigma}_{u\xi^*} + \tilde{\beta} \tilde{\kappa} s_{22} \quad (21)$$

$$s_{11} = \sigma_u^2 + \tilde{\beta} (2\tilde{\sigma}_{u\xi^*} + \tilde{\beta} \tilde{\kappa} s_{22}). \quad (22)$$

In essence, we have transformed a problem with non-classical measurement error into an equivalent problem with classical measurement error but different parameter values. In the transformed system, the extent of measurement error is controlled by $\tilde{\kappa}$ and regressor endogeneity is controlled by $\tilde{\sigma}_{u\xi^*}$. Instrument invalidity is controlled by the *same* parameter in both the original and transformed parameterizations: $\sigma_{u\zeta}$. While $\tilde{\kappa}$ is scale-free, $\sigma_{u\zeta}$ and $\tilde{\sigma}_{u\xi^*}$ are not. For this reason, when we derive the restrictions implied by (18)–(22) below we

will express them in terms of correlations rather than covariances, namely

$$\rho_{u\zeta} \equiv \text{Cor}(\zeta, u), \quad \rho_{u\xi^*} \equiv \text{Cor}(u, \xi^*). \quad (23)$$

Note that

$$\rho_{u\xi^*} = \frac{\sigma_{u\xi^*}}{\sigma_u \sqrt{\tilde{\kappa} S_{22}}} = \frac{(1 + \psi)\sigma_{u\xi^*}}{\sigma_u \sqrt{(1 + \psi)^2 \kappa S_{22}}} = \frac{\tilde{\sigma}_{u\xi^*}}{\sigma_u \sqrt{\tilde{\kappa} S_{22}}} \quad (24)$$

so that $\rho_{u\xi^*}$, unlike $\sigma_{u\xi^*}$, is unaffected by the re-parameterization in (18)–(22). In summary, we can proceed *as though* the measurement error were classical by working in terms of $(\rho_{u\zeta}, \rho_{u\xi^*}, \tilde{\kappa})$. Any restrictions on ψ , for example in the case of a binary T^* , can be addressed in a second step. In the following section, we derive the joint restrictions between these parameters and the identified set for β .

2.3 Joint Restrictions

A key point of this paper is that beliefs over measurement error, regressor endogeneity, and instrument invalidity are mutually constrained by each other and the data. The following result makes this intuition precise by expressing $\rho_{u\zeta}$ as an explicit function of $\rho_{u\xi^*}$ and $\tilde{\kappa}$, given particular values of the reduced form correlations.

Proposition 2.1. *Under Assumptions 2.1–2.3,*

$$\rho_{u\zeta} = \frac{r_{23}\rho_{u\xi^*}}{\tilde{\kappa}^{1/2}} - (r_{12}r_{23} - r_{13}\tilde{\kappa}) \left[\frac{1 - \rho_{u\xi^*}^2}{\tilde{\kappa}(\tilde{\kappa} - r_{12}^2)} \right]^{1/2} \quad (25)$$

where $r_{12} \equiv \text{Cor}(\varepsilon, \xi)$, $r_{13} \equiv \text{Cor}(\varepsilon, \zeta)$, and $r_{23} \equiv \text{Cor}(\xi, \zeta)$.

Equation 25 is the first ingredient in our characterization of the joint restrictions between measurement error, regressor endogeneity, and instrument invalidity. The second is a bound on $\tilde{\kappa}$ that limits the possible extent of measurement error in the data.

Proposition 2.2. *Under Assumptions 2.1–2.3, $\tilde{\kappa} \in (L, 1]$ where*

$$L \equiv \frac{r_{12}^2 + r_{23}^2 - 2r_{12}r_{23}r_{13}}{1 - r_{13}^2} > \max \{r_{12}^2, r_{23}^2\}, \quad (26)$$

and the reduced-form correlations r_{12}, r_{23} , and r_{13} are as defined in Proposition 2.1.

Because it places a lower bound on $\tilde{\kappa}$, namely L , Proposition 2.2 places an *upper bound* on the extent of measurement error. The derivation of this bound relies on two simpler but weaker bounds. The first, $\tilde{\kappa} > r_{12}^2$, corresponds to the familiar “reverse regression bound”

under classical measurement error. The second, $\tilde{\kappa} > r_{23}^2$, is in essence a reverse regression bound constructed from the IV *first-stage*. The bound $\tilde{\kappa} > L$ is strictly tighter than both of these bounds, as it incorporates information from all three of the reduced form correlations: r_{12}, r_{23} , and r_{13} . [Proposition 2.2](#) does not, however, allow us to rule out the possibility that there is no measurement error: $\tilde{\kappa} = 1$ always satisfies the bounds regardless of the values of the reduced form correlations.

Together, [Proposition 2.1](#) and [Proposition 2.2](#) provide joint restrictions on instrument invalidity, regressor endogeneity, and measurement error. In particular, the reduced form covariance matrix Σ both bounds $\tilde{\kappa}$ and gives $\rho_{u\zeta}$ as an explicit function of $\rho_{u\xi^*}$ and $\tilde{\kappa}$. These restrictions in fact constitute the sharp identified set, as we now show.

Theorem 2.1. *Suppose that T^* has full support, Σ is finite and positive definite with $s_{23} \neq 0$, and $(\varphi_y, \varphi_T, \varphi_z)$ are likewise finite. Under Assumptions [2.1–2.3](#), the restrictions $\psi > -1$, $|\rho_{u\xi^*}| < 1$, $\tilde{\kappa} \in (L, 1]$, and [\(25\)](#) characterize the sharp identified set for $(\rho_{u\zeta}, \rho_{u\xi^*}, \tilde{\kappa}, \psi, \tau)$.*

The additional assumption $s_{23} \neq 0$ in [Theorem 2.1](#) is a reduced form version of the structural instrument relevance condition from [Assumption 2.1](#) (iii); it requires that z is correlated with T even after projecting out \mathbf{x} . Note that [Theorem 2.1](#) imposes no cross-restrictions between the parameters $\tilde{\kappa}$, ψ , τ , and $\rho_{u\xi^*}$. In contrast, $\rho_{u\zeta}$ is completely determined by $\tilde{\kappa}$ and $\rho_{u\xi^*}$ by [\(25\)](#). Moreover, ψ , τ and $\rho_{u\xi^*}$, unlike $\tilde{\kappa}$, are completely unrestricted by observables. As shown in the following result, our assumptions also bound the instrument invalidity parameter $\rho_{u\zeta}$, despite placing no restriction on regressor endogeneity.

Corollary 2.1. *Under the conditions of [Theorem 2.1](#), $\rho_{u\zeta}$ has a non-trivial one-sided bound. If $r_{12}r_{23} < Lr_{13}$, then $\rho_{u\zeta} \in (-|r_{23}|/\sqrt{L}, 1)$; otherwise $\rho_{u\zeta} \in (-1, |r_{23}|/\sqrt{L})$, where L is defined in [Proposition 2.2](#). These bounds are sharp.*

Because $L > r_{23}^2$, [Corollary 2.1](#) always rules out a range of values for $\rho_{u\zeta}$. Notice, however, that it never rules out $\rho_{u\zeta} = 0$. This is unsurprising given that it is known to be impossible to test for instrument validity in the model we consider here. Unfortunately, and also unsurprisingly, the model itself places no restrictions on the causal effect β .

Corollary 2.2. *Under the conditions of [Theorem 2.1](#), the sharp identified set for the causal effect of interest, β , is $(-\infty, \infty)$.*

The only way to learn about β in this model is to impose beliefs. In our examples below we consider simple interval restrictions on $\tilde{\kappa}$ and $\rho_{u\xi^*}$. [Proposition B.1](#) in the appendix shows how interval restrictions on $\tilde{\kappa}$ and $\rho_{u\xi^*}$ tighten the bounds for $\rho_{u\zeta}$ from [Corollary 2.1](#). [Proposition B.2](#) shows that any restriction on $\rho_{u\xi^*}$ that rules out values arbitrarily close to

-1 or 1 yields finite bounds for $\tilde{\beta}$. In the case of classical measurement error, $\psi = 0$ and hence bounds for $\tilde{\beta}$ are equivalent to bounds for β . In the general case, translating bounds for $\tilde{\beta}$ into bounds for β requires restrictions on ψ . When T^* is binary, the data provide such restrictions. In the following section we derive these restrictions and show how to incorporate them into our partial identification exercise.

3 The Case of a Binary T^*

In many applied studies the regressor of interest is binary: $T^*, T \in \{0, 1\}$. In this case [Theorem 2.1](#) no longer applies: the data impose additional restrictions on ψ through the support restriction on T^* . We now show how to extend our analysis from [section 2](#) to incorporate the additional information available in the binary T^* case. Similar reasoning can be applied when T^* has an arbitrary discrete support set, although we do not pursue the general case here. To begin, we define some additional notation specific to the binary setting. First let $p^* \equiv \mathbb{P}(T^* = 1)$ and $p \equiv \mathbb{P}(T = 1)$. Next define the mis-classification error rates α_0 and α_1 as follows:

$$\alpha_0 \equiv \mathbb{P}(T = 1 | T^* = 0), \quad \alpha_1 \equiv \mathbb{P}(T = 0 | T^* = 1). \quad (27)$$

The parameter α_0 equals the probability of an *upwards* mis-classification error, observing $T = 1$ when $T^* = 0$. In contrast, α_1 equals the probability of a *downwards* mis-classification error, observing $T = 0$ when $T^* = 1$. Using this notation, we can express ψ, τ and w as functions of (α_0, α_1) as follows.

Lemma 3.1. *Suppose that $T^*, T \in \{0, 1\}$ and define (α_0, α_1) as in (27). Then*

- (i) $\psi = -(\alpha_0 + \alpha_1)$
- (ii) $\tau = \alpha_0$
- (iii) $w = (T - \alpha_0) - (1 - \alpha_0 - \alpha_1)T^*$.

[Lemma 3.1](#) reveals two important features of the binary T^* case. First, while ψ could be positive or negative in the general case, it must be *negative* in the binary case. Second, while τ and ψ are in general two free parameters, they are linked through their joint dependence on α_0 in the binary case. Under [Assumption 2.1](#) (v), we have $\psi > -1$. By [Lemma 3.1](#) this is equivalent to $\alpha_0 + \alpha_1 < 1$ when T^* is binary. The following Lemma exploits this fact to relate p^* to p and to yield a simple expression for σ_w^2 in terms of (α_0, α_1) and p .

Lemma 3.2. *Suppose that $T^*, T \in \{0, 1\}$ and define (α_0, α_1) as in (27). Then, provided that $\alpha_0 + \alpha_1 \neq 1$,*

$$(i) \quad p^* = (p - \alpha_0)/(1 - \alpha_0 - \alpha_1)$$

$$(ii) \quad \sigma_w^2 = \alpha_1(1 - \alpha_0) + (1 - p)(\alpha_0 - \alpha_1)$$

We now have two equations for σ_w^2 in the binary T^* case: (21), and Lemma 3.2 (ii). Equating these yields the following cross-restriction between ψ and $\tilde{\kappa}$.

Proposition 3.1. *Let $T^*, T \in \{0, 1\}$ and suppose that Σ is positive definite. Then under Assumptions 2.1–2.3, $\underline{\psi}(\tilde{\kappa}) \leq \psi \leq \bar{\psi}(\tilde{\kappa})$ where*

$$\bar{\psi}(\tilde{\kappa}) \equiv \frac{-s_{22}(1 - \tilde{\kappa})}{\max\{p, 1 - p\}}, \quad \underline{\psi}(\tilde{\kappa}) \equiv \begin{cases} \frac{-s_{22}(1 - \tilde{\kappa})}{\min\{p, 1 - p\}}, & s_{22}(1 - \tilde{\kappa}) \leq m(p) \\ 2\sqrt{p(1 - p) - s_{22}(1 - \tilde{\kappa})} - 1, & s_{22}(1 - \tilde{\kappa}) > m(p) \end{cases}$$

with $m(p) \equiv \max\{(1 - p)(2p - 1), p(1 - 2p)\}$ and $p \equiv \mathbb{P}(T = 1)$.

The intuition behind Proposition 3.1 is as follows. In the binary T^* case, both $\tilde{\kappa}$ and ψ are functions of the mis-classification probabilities α_0 and α_1 . By definition these must lie between zero and one, and by Assumption 2.1 (v) they also satisfy $\alpha_0 + \alpha_1 < 1$. This region is depicted in Figure 1. Since $\sigma_w^2 = s_{22}(1 - \tilde{\kappa})$ by (21), choosing a value for $\tilde{\kappa}$ is equivalent to choosing a value of σ_w^2 . Hence, solving the expression from Lemma 3.2 (ii), the choice of $\tilde{\kappa}$ determines α_1 as a function of α_0 . The figure depicts three such functions, corresponding to three different choices of $\tilde{\kappa}$: $L < \tilde{\kappa}_1 < \tilde{\kappa}_2$. Since $L < \tilde{\kappa}$ by Proposition 3.1, the first of these choices gives the *outer envelope* of this family of functions. The bounds for ψ are determined by first pinning down a single function from this family by choosing a feasible value of $\tilde{\kappa}$, and then finding all values of C such that $\alpha_0 + \alpha_1 = C$ intersects this function. The minimum value of $(\alpha_0 + \alpha_1)$ always occurs at a corner. In the figure we set $p > 1/2$ so that the minimum occurs at $s_{22}(1 - \tilde{\kappa})/p$. The maximum, indicated by the filled circles in the figure, can either be interior (red) or occur at a corner (blue). A corner maximum occurs when $\tilde{\kappa}$ is sufficiently large, or equivalently σ_w^2 is sufficiently small. Finally, Lemma 3.1 converts bounds for $(\alpha_0 + \alpha_1)$ into bounds for ψ .

In some cases, additional a priori information may be available to further restrict α_0 and α_1 and hence $(\tilde{\kappa}, \psi)$. For example, under *one-sided* mis-classification, either α_0 or α_1 is known to be zero. Another such case is that of *symmetric* mis-classification, in which $\alpha_0 = \alpha_1$. A third example concerns settings in which auxiliary data suggest that $p^* \approx p$. This corresponds to the restriction $\alpha_1 \approx \alpha_0(1 - p)/p$. Each of these three special cases

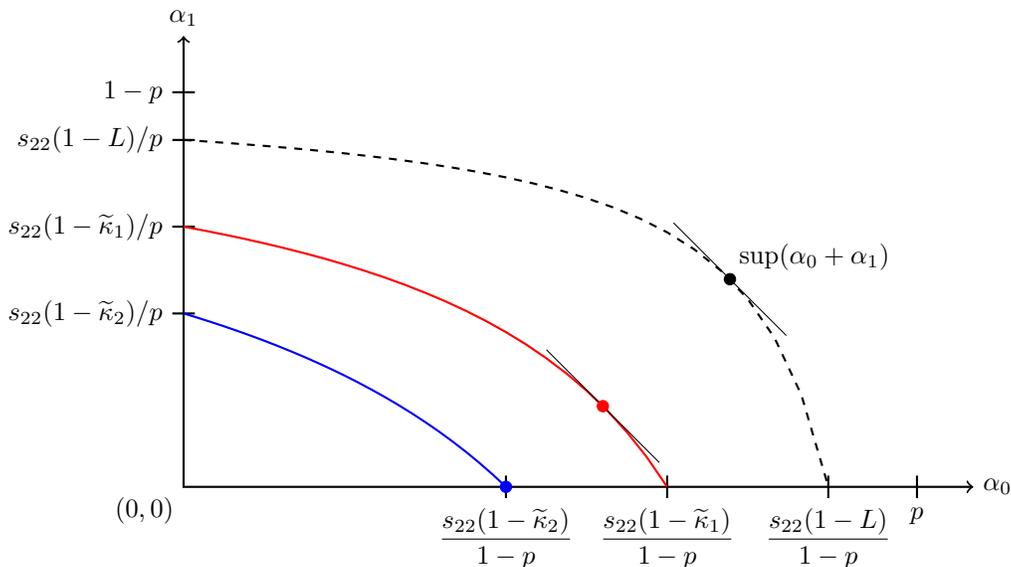


Figure 1: Restrictions on α_0 and α_1 for three values of $\tilde{\kappa}$: $L < \tilde{\kappa}_1 < \tilde{\kappa}_2$ where L is as defined in [Proposition 2.2](#). Here $p > 1/2$ so the minimum value of $(\alpha_0 + \alpha_1)$ for a fixed $\tilde{\kappa}$ occurs at $s_{22}(1 - \tilde{\kappa})/p$. The maximum value of $(\alpha_0 + \alpha_1)$ is interior for $\tilde{\kappa}$ sufficiently small (L and $\tilde{\kappa}_2$) and occurs at a corner for $\tilde{\kappa}$ sufficiently large ($\tilde{\kappa}_2$). Here the corner solution has $\alpha_0 = 0$ since $p > 1/2$. The supremum of $(\alpha + \alpha_1)$ occurs at $\tilde{\kappa} = L$ and the minimum at $\tilde{\kappa} = 0$, i.e. zero measurement error.

yields a linear equality restriction of the form $M_0\alpha_0 + M_1\alpha_1 = 0$ and reduces the number of unknown parameters by one. Geometrically this takes the form of a line with non-negative slope passing through the origin of [Figure 1](#), meaning that ψ is an explicit function of $\tilde{\kappa}$. In the case of symmetric mis-classification, for example, ψ is determined by the intersection of the 45-degree line and the curve corresponding to a given choice of $\tilde{\kappa}$.

Without support restrictions, we know from [Theorem 2.1](#) that the data are uninformative about ψ . [Proposition 3.1](#) shows that when the support of T^* is restricted to $\{0, 1\}$ this is no longer the case: the observables restrict ψ , and $\tilde{\kappa}$ and ψ are mutually constrained. [Proposition B.3](#) in the Appendix shows how to use these restrictions to bound β . To summarize, the logic of [Proposition B.2](#) shows that $\tilde{\beta}$ is bounded so long as $\rho_{u\xi^*}$ is restricted *a priori* to lie in a strict subset of $(-1, 1)$. [Proposition B.3](#) combines this observation with [Proposition 3.1](#) to yield bounds for β via [\(18\)](#).

As we show in our empirical example from [subsection 5.3](#) below, the restrictions imposed by [Proposition 3.1](#), in concert with [Proposition 2.1](#) and [Proposition 2.2](#), can be very informative in practice. Moreover, they allow us to treat the continuous and binary T^* cases within a common, regression-based framework. However, these restrictions do not necessarily constitute the sharp identified set when T^* is binary. For example, knowledge of the conditional distribution of $T|\mathbf{x}$ could in principle provide further restrictions on (α_0, α_1) . Exploiting this

information, however, would require modeling objects over which applied researchers remain agnostic when reporting OLS and IV regressions, even with a binary T^* . Accordingly we do not pursue this possibility further here.⁸

4 Elicitation and Inference

We now describe how to use our results from above to carry out Bayesian inference. We present two approaches: inference for the identified set Θ and inference for the partially identified parameter θ . We focus throughout on two cases that are common in applications: first a regressor T^* without support restrictions that is subject to classical measurement error, and second a binary T^* as examined in [section 3](#) above. [Sections 4.1](#) and [4.2](#) consider classical measurement error, i.e. $\psi = 0$, in which case $\tilde{\beta} = \beta$, $\tilde{\kappa} = \kappa$, etc. [Section 4.3](#) explains the differences that arise when T^* is binary. [Online Appendix D](#) provides some discussion of the relationship between Bayesian and Frequentist inference in partially identified models.

Our approach relies on the principle that the choice of parameterization should make clear how any prior beliefs that cannot be falsified by data affect the ultimate result. For this reason, our derivations from above relate the identified set Θ for the structural parameters θ to the reduced form parameters $\varphi \equiv (\Sigma, \varphi_y, \varphi_T, \varphi_z)$, i.e. $\Theta(\varphi)$, such that any inferences we draw about θ depend on the data only through φ .⁹ Because φ is point-identified, inference for this parameter vector is standard. We begin by assuming that the researcher has computed a posterior for φ . [Section subsection 4.4](#) discusses how to obtain one.

We elicit researcher beliefs in the form of sign and interval restrictions, \mathcal{R} , over regressor endogeneity, instrument invalidity, and measurement error. Intersecting $\Theta(\varphi)$ with \mathcal{R} adds relatively weak prior information to restrict the identified set in a transparent manner. To simplify the elicitation of \mathcal{R} , our results in [section 2](#) are expressed in terms of scale-free parameters. The regressor endogeneity parameter $\rho_{u\xi^*}$ and the instrument invalidity parameter $\rho_{u\zeta}$ are correlations, and have the same meaning regardless of whether the measurement error is classical or non-classical. In practice, a researcher might state a sign restriction for one or both of these quantities, along with an upper bound that is thought to represent an implausibly large extent of correlation. The appropriate way to elicit information about measurement error depends on the nature of that error. In the classical measurement error case $\psi = 0$ and hence $\kappa = \tilde{\kappa}$. In this case, one could elicit interval restrictions over the scale-free variance ratio κ . Because κ is defined net of covariates \mathbf{x} , it may be easier in some settings

⁸For related results, see [DiTraglia and García-Jimeno \(2019\)](#) who derive the sharp identified set for a mis-classified, binary endogenous regressor given a valid instrument with discrete support, in an additively separable model with arbitrary dependence on exogenous covariates.

⁹This is called a *transparent parameterization* in the statistics literature: see, e.g., [Gustafson \(2015\)](#).

to instead elicit $\lambda \equiv \text{Var}(T^*)/\text{Var}(T)$ and transform this to κ via $\kappa = (\lambda - R_{T,\mathbf{x}}^2)/(1 - R_{T,\mathbf{x}}^2)$ where $R_{T,\mathbf{x}}^2$ is the R-squared from a regression of T on \mathbf{x} . In the binary T^* case, neither ψ nor $\tilde{\kappa}$ is a natural parameter over which to elicit beliefs, but both are completely determined by α_0 and α_1 . It is over these mis-classification probabilities, also scale-free, that researchers would most likely be able to state beliefs.

4.1 Inference for the Identified Set

We first consider Bayesian posterior inference for the identified set for θ rather than the structural parameter vector itself. If $\varphi^{(j)}$ is a draw from the posterior for φ , then $\Theta(\varphi^{(j)}) \cap \mathcal{R}$ is a draw from the posterior distribution for the identified set for θ under researcher beliefs \mathcal{R} . By collecting a large number of these draws, one can summarize the posterior in a variety of different ways. First, one can construct a credible interval for the identified set of a particular structural parameter, such as $\rho_{u\zeta}$ or β , under a set of *a priori* restrictions \mathcal{R} . If \mathcal{R} restricts $\rho_{u\xi}^*$ to a proper subset of $(-1, 1)$, then [Proposition B.1](#) yields two sided bounds for the instrument invalidity parameter $\rho_{u\zeta}$, while [Proposition B.2](#) yields two-sided bounds for the causal effect β . Suppose we wish to form a 90% credible interval for the identified set \mathcal{B} for β . To construct this interval, start with the conditional identified set $\mathcal{B}(\bar{\varphi})$ evaluated at the posterior mean $\bar{\varphi}$ and expand this interval outwards symmetrically until the resulting interval contains 90% of the identified sets. As we show in our empirical examples below, such intervals for β can in some cases be surprisingly informative, despite relaxing the requirement that z is a valid instrument.

Second, one can use the posterior to quantify the extent to which a particular set of *a priori* researcher beliefs \mathcal{R} accords with the data by calculating the posterior probability that the intersection of $\Theta(\varphi)$ with \mathcal{R} is empty. Consider, for example, a researcher who believes that selection is negative ($\rho_{u\xi^*} < 0$) and wishes to assess whether this is compatible with a belief that her instrument is valid ($\rho_{u\zeta} = 0$). If we define \mathcal{R} to be the intersection of these two restrictions, then calculating the fraction of sets $\Theta(\varphi^{(j)}) \cap \mathcal{R}$ that are nonempty yields the posterior probability that we cannot rule out instrument validity under a particular assumption about the direction of selection. We abbreviate this as $\mathbb{P}(\text{Valid})$ in our empirical examples below. If $\mathbb{P}(\text{Valid})$ is small, the data strongly suggest that the assumed direction of selection is incompatible with instrument validity. More generally, consider any restriction \mathcal{R} . Calculating the fraction of sets $\Theta(\varphi^{(j)}) \cap \mathcal{R}$ that are empty gives the posterior probability that \mathcal{R} can be ruled out, a probability that we abbreviate as $\mathbb{P}(\emptyset)$ in our empirical examples below. If $\mathbb{P}(\emptyset)$ is small but nonzero, a researcher who feels confident in her *a priori* beliefs could elect to discard the draws $\varphi^{(j)}$ for which $\Theta(\varphi^{(j)}) \cap \mathcal{R}$ is empty. If $\mathbb{P}(\emptyset)$ is large, this suggests

that the beliefs encoded in \mathcal{R} are suspect, given the data. When \mathcal{R} restricts two or more dimensions of $(\rho_{u\zeta}, \rho_{u\xi^*}, \kappa)$, a large value of $\mathbb{P}(\emptyset)$ indicates that the corresponding researcher beliefs are mutually incompatible *a posteriori*. This exercise illustrates an important general point of our approach. By making explicit the relationship between measurement error, treatment endogeneity, and instrument invalidity, our method allows researchers to learn whether their beliefs over these different dimensions of the problem cohere.

4.2 Inference for the Partially Identified Parameter

Our second approach makes posterior probability statements about the partially identified parameter θ , by averaging both over reduced form draws $\varphi^{(j)}$ and a conditional prior placed on $\Theta(\varphi^{(j)})$. Carrying out inference for θ rather than its identified set is attractive. For example, it allows one to compute the posterior probability that β is positive. This, however, comes at a cost: the need to specify a conditional prior over the identified set. Because it may be difficult in practice to elicit a fully informative prior, following [Moon and Schorfheide \(2012\)](#) we recommend placing a uniform reference prior on $\Theta(\varphi^{(j)}) \cap \mathcal{R}$ (see [Appendix C](#) for implementation details). Our use of this prior is intended to represent prior ignorance over $\Theta(\varphi^{(j)}) \cap \mathcal{R}$. Unavoidably, uniformity in one parameterization could imply a highly informative prior in some different parameterization. We emphasize, however, that the uniform serves here as a reference prior only. As such, one need not take it completely literally but could instead consider, for example, what kinds of deviations from uniformity would be necessary to support a particular belief about β .

A prior on the conditional identified set cannot be updated by the data. As such its influence on the posterior does not vanish as the sample size grows. For this reason, some caution is warranted when carrying out posterior inference for θ . A researcher who is concerned about this issue may wish to carry out a Bayesian robustness exercise over a class of priors supported on the conditional identified set. If this class includes all possible priors over $\Theta(\varphi^{(j)}) \cap \mathcal{R}$, the resulting bounds on posterior probabilities for θ will coincide with our inferences for the identified set from [4.1](#). While robust, such inferences are inherently conservative, as they summarize only the most extreme points of $\Theta(\varphi^{(j)}) \cap \mathcal{R}$. Suppose for example that each draw $\Theta(\varphi^{(j)}) \cap \mathcal{R}$ includes a single point that implies a negative value of β . Then, inference for the identified set \mathcal{B} would produce no evidence against the claim that $\beta \leq 0$. In contrast, *any* reasonable prior over $\Theta(\varphi^{(j)}) \cap \mathcal{R}$, such as our uniform reference prior, would give 100% posterior probability to $\{\beta > 0\}$.

4.3 The Binary T^* Case

We now summarize the modifications to our inference approaches from [subsection 4.1](#) and [subsection 4.2](#) that are required to treat the binary T^* case from [section 3](#). In this case, ψ is in general non-zero and hence $\tilde{\kappa}$ and $\tilde{\beta}$ need not equal κ and β . Note, however, that the meaning of $\rho_{u\zeta}$, along with that of $\rho_{u\xi^*}$, is unchanged in the binary T^* case. Moreover, [\(25\)](#) does not involve ψ , nor does [Proposition 3.1](#) impose cross-restrictions between $\rho_{u\zeta}$ and ψ . As such, to carry out inference for $\rho_{u\zeta}$ we can proceed *exactly* as we did in the classical measurement error case: all that changes is the interpretation of $\tilde{\kappa}$. This underscores a key advantage of working with a scale-free parameterization: the interpretations of $\rho_{u\zeta}$ and $\rho_{u\xi^*}$ do not depend on ψ . [Proposition 3.1](#) does, however, create a cross-restriction between $\tilde{\kappa}$ and ψ . If $\tilde{\beta}$ were our parameter of interest, we could ignore this fact and proceed as though the measurement error were classical. Because we are actually interested in $\beta = (1 + \psi)\tilde{\beta}$, an extra step is required. To carry out inference for the identified set for β , we rely on [Proposition B.3](#) to yield bounds for β at any given reduced form draw $\boldsymbol{\varphi}^{(j)}$. To carry out inference for the partially identified *parameter* β , we first draw $\boldsymbol{\varphi}^{(j)}$ and then sample $(\tilde{\kappa}^{(j)}, \rho_{u\xi^*}^{(j)}, \rho_{u\zeta}^{(j)})$ uniformly on the resulting conditional identified set, as described in [subsection 4.2](#). We then draw $\psi^{(j)}$ uniformly from the interval $[\underline{\psi}(\tilde{\kappa}^{(j)}), \overline{\psi}(\tilde{\kappa}^{(j)})]$ defined in [Proposition 3.1](#). Given these draws, we construct the implied draw for $\beta^{(j)}$ using the derivations from [section 2](#).

4.4 Posterior Inference for the Reduced Form Parameters

To implement the procedures from [subsection 4.1](#) and [subsection 4.2](#) the researcher must first obtain a posterior for the reduced form parameters. As we showed above in [section 2](#), the reduced form regression slopes $(\boldsymbol{\varphi}_y, \boldsymbol{\varphi}_T, \boldsymbol{\varphi}_z)$ play no role in determining the identified set for $\boldsymbol{\theta}$. For this reason, we only require posterior draws for Σ . In our empirical examples below, we adopt the following simple approach. Given an iid sample of n observations $(y_i, T_i, z_i, \mathbf{x}_i)$, let $\mathbf{y} = (y_1, \dots, y_n)'$ and define \mathbf{T} and \mathbf{z} analogously. Further define $X' = (\mathbf{x}'_1, \dots, \mathbf{x}'_n)$ and $Y = [\mathbf{y} \quad \mathbf{T} \quad \mathbf{z}]$. We draw Σ from an Inverse-Wishart(ν, S) distribution where

$$\nu = n - k + 3 + 1, \quad S = (Y - X\hat{B})'(Y - X\hat{B}), \quad \hat{B} = (X'X)^{-1}X'Y$$

and k is the dimension of the exogenous covariate vector \mathbf{x}_i . Note that the mean of this distribution equals $S/(n-k)$, the sample covariance matrix of OLS residuals from the reduced form regressions given in [\(6\)](#). The Inverse-Wishart(ν, S) distribution is the marginal posterior for Σ in the multivariate reduced form regression obtained by stacking [\(6\)](#) under a Jeffreys prior and normal errors (see e.g. [Zellner, 1971](#), Section 8.1).

For simplicity, we draw the reduced form covariance matrix from an Inverse-Wishart posterior in both the classical measurement error and binary T^* cases. Of course, the reduced form errors cannot be normal if any of the variables (y, T, z) is discrete. Nonetheless, our Inverse-Wishart posterior for Σ is still centered at $S/(n - k)$ and is approximately normal in large samples under mild conditions, as we discuss in Appendix D. Note that the bounds for ψ from Proposition 3.1 in the binary T^* case involve p . To address this minor complication, we adopt an empirical Bayes approach, setting p equal to the sample analogue \hat{p} . Because this quantity is very precisely estimated, its effect on our inferences is negligible. An alternative to our Inverse-Wishart posterior for Σ is the Bayesian Bootstrap approach followed by Bollinger and van Hasselt (2017).

5 Empirical Examples

We now present three empirical examples illustrating how the framework described above can be applied in practice. The examples in Sections 5.1 and 5.2 involve a continuous treatment which we assume is subject to classical measurement error, i.e. $\psi = 0$, $\tilde{\kappa} = \kappa$ and $\tilde{\beta} = \beta$. In contrast, the example in Section 5.3 involves a binary treatment, so that any measurement error that is present must be non-classical.

5.1 The Colonial Origins of Comparative Development

Acemoglu et al. (2001) study the effect of institutions on GDP per capita using a cross-section of 64 countries. Because institutional quality is endogenous, they use differences in the mortality rates of early western settlers across colonies as an instrumental variable. We consider their benchmark specification

$$\begin{aligned} \log \text{GDP/capita} &= \text{constant} + \beta (\text{Institutions}) + u \\ \text{Institutions} &= \text{constant} + \pi (\log \text{Settler Mortality}) + v \end{aligned}$$

which does not include covariates.¹⁰ This yields an IV estimate of 0.94 with a standard error of 0.16 – nearly twice as large as the corresponding OLS estimate of 0.52 with a standard error of 0.06. The authors attribute this disparity to classical measurement error:

This estimate is highly significant ...and in fact larger than the OLS estimates ...This suggests that measurement error in the institutions variables that creates attenuation

¹⁰Additional results, available upon request, consider alternative specifications that include covariates. The results are essentially unchanged.

bias is likely to be more important than reverse causality and omitted variables biases.
 (Acemoglu et al., 2001, p. 1385)

Acemoglu et al. (2001) state two beliefs that are relevant for our partial identification exercise. First, their discussion implies there is likely a positive correlation between “true” institutions and the main equation error term u . This could arise from reverse causality – wealthier societies can afford better institutions – or omitted variables, such as legal origin or British culture, which are likely to be positively correlated with present-day institutional quality. We encode this belief using the prior restriction $0 < \rho_{u\xi^*} < 0.9$ below, ruling out only unreasonably large values of treatment endogeneity.¹¹ Second, in a footnote that uses an alternative measure of institutions as an instrument for the first, the authors argue that measurement error could be substantial.¹² Taken at face value, the calculations from this footnote imply a point estimate of $\kappa = 0.6$ which would mean that 40 percent of the variation in measured institutions is noise.¹³ Below we consider two alternative ways of encoding this auxiliary information about κ .

Results for the Colonial Origins example appear in Table 1. Estimates and bounds for β indicate the percentage increase in GDP per capita that would result from a one point increase in the quality of institutions, as measured by average protection against expropriation risk. All other values in the table are unitless: they are either probabilities, correlations, or variance ratios. OLS and IV estimates and standard errors, along with an estimate of the lower bound L for κ , appear in the first row of Panel (I). Panel (II) presents inferences for the identified set. The first column of Panel (II) gives the fraction of posterior draws for the reduced form parameters that yield an empty identified set, while the second column gives the fraction that are compatible with a valid instrument: $\rho_{u\zeta} = 0$. The third and fourth columns of Panel (II) present 90% posterior credible intervals for the identified sets for $\rho_{u\zeta}$ and β , constructed by symmetrically expanding around the conditional identified set evaluated at the posterior mean for Σ , as described in subsection 4.1. In contrast, panel (III) presents posterior medians and 90% highest posterior density intervals for $\rho_{u\zeta}$ and β , based on the uniform reference prior described in subsection 4.2.

¹¹By Corollary 2.2, the identified set for β is $(-\infty, \infty)$ unless $\rho_{u\xi^*}$ is restricted. Here we impose the researchers’ stated belief that $\rho_{u\xi^*} > 0$ along with an extremely conservative upper bound for $\rho_{u\xi^*}$ of 0.9.

¹²Footnote #19 of Acemoglu et al. (2001) states “We can ascertain, to some degree, whether the difference between OLS and 2SLS estimates could be due to measurement error by making use of an alternative measure of institutions ...This suggests that ‘measurement error’ in the institutions variables ...is of the right order of magnitude to explain the difference between the OLS and 2SLS estimates.”

¹³Suppose T_1 and T_2 are two measures of institutions that are subject to classical measurement error: $T_1 = T^* + w_1$ and $T_2 = T^* + w_2$. Both T_1 and T_2 suffer from precisely the same degree of endogeneity, because they inherit this problem from T^* alone under the assumption of classical measurement error. Thus, the OLS estimator based on T_1 converges to $\kappa(\beta + \sigma_{T^*u}/\sigma_{T_1}^2)$ while the IV estimator that uses T_2 to instrument for T_1 converges to $\beta + \sigma_{T^*u}/\sigma_{T_2}^2$. The ratio identifies κ : $0.52/0.87 \approx 0.6$.

	(I) Summary Statistics			(II) Inference for Θ			(III) Inference for θ		
	OLS	IV	L	$\mathbb{P}(\emptyset)$	$\mathbb{P}(\text{Valid})$	$\rho_{u\zeta}$	β	$\rho_{u\zeta}$	β
Colonial Origins ($n = 64$)	0.52 (0.06)	0.94 (0.16)	0.54						
$(\kappa, \rho_{u\xi^*}) \in (0, 0.6] \times [0, 0.9]$				0.26	-	$[-, -]$	$[-, -]$	-	$[-, -]$
$(\kappa, \rho_{u\xi^*}) \in (0.6, 1] \times [0, 0.9]$				0.00	0.30	$[-1.00, 0.61]$	$[-0.66, 1.05]$	-0.57	0.49
								$[-0.81, -0.17]$	$[0.01, 0.94]$

Table 1: Results for Colonial Origins Example. Panel (I) contains OLS and IV estimates and standard errors, and the posterior mean estimate for the lower bound L for κ and $\rho_{u\xi^*}$. Panel (II) gives posterior inference for the identified set. The column $\mathbb{P}(\emptyset)$ gives the fraction of reduced form parameter draws that yield an empty identified set, while $\mathbb{P}(\text{Valid})$ gives the fraction of reduced form parameter draws compatible with a valid instrument ($\rho_{u\zeta} = 0$). The remaining columns of Panel (II) give 90% posterior credible intervals for the identified set for $\rho_{u\zeta}$ and β . In contrast, Panel (III) presents posterior medians and 90% credible intervals for the partially identified parameters $\rho_{u\zeta}$ and β under a conditionally uniform reference prior. See [section 4](#) for details.

	(I) Summary Statistics			(II) Inference for Θ			(III) Inference for θ		
	OLS	IV	L	$\mathbb{P}(\emptyset)$	$\mathbb{P}(\text{Valid})$	$\rho_{u\zeta}$	β	$\rho_{u\zeta}$	β
Was Weber Wrong? ($n = 452$)	0.10 (0.01)	0.19 (0.03)	0.49						
$(\kappa, \rho_{u\xi^*}) \in (0, 1] \times [-0.9, 0]$				0.00	1.00	$[-0.24, 0.57]$	$[-0.02, 1.00]$	0.32	0.37
$(\kappa, \rho_{u\xi^*}) \in (0.8, 1] \times [-0.9, 0]$				0.00	1.00	$[-0.24, 0.45]$	$[0.06, 0.66]$	$[-0.10, 0.82]$	$[0.12, 0.61]$
								0.06	0.22
								$[-0.15, 0.28]$	$[0.10, 0.42]$

Table 2: Results for “Was Weber Wrong?” (Section 5.2). Panel (I) contains OLS and IV estimates and standard errors, and the posterior mean estimate for the lower bound L for κ and $\rho_{u\xi^*}$. Panel (II) gives posterior inference for the identified set. The column $\mathbb{P}(\emptyset)$ gives the fraction of reduced form parameter draws that yield an empty identified set, while $\mathbb{P}(\text{Valid})$ gives the fraction of reduced form parameter draws compatible with a valid instrument ($\rho_{u\zeta} = 0$). The remaining columns of Panel (II) give 90% posterior credible intervals for the identified set for $\rho_{u\zeta}$ and β . In contrast, Panel (III) presents posterior medians and 90% credible intervals for the partially identified parameters $\rho_{u\zeta}$ and β under a conditionally uniform reference prior. See [section 4](#) for details.

We first consider an *a priori* restriction that $\kappa < 0.6$, placing a *lower bound* on the extent of measurement error. This restriction comes from personal communication with one of the authors of [Acemoglu et al. \(2001\)](#).¹⁴ Under this restriction, approximately 26 percent of the draws for the reduced form parameters yield an *empty* identified set, as shown in the first column of Panel (II). Intuitively, this means that there are covariance matrices Σ that are close to the maximum likelihood estimate $\widehat{\Sigma}$ but which rule out the region $(\kappa, \rho_{u\xi^*}) \in (0, 0.6] \times [0, 0.9]$. The problem is not the restriction on $\rho_{u\xi^*}$ but on κ : the data place no restrictions on the extent of treatment endogeneity although they do provide an upper bound on the extent of measurement error, as shown in [Theorem 2.1](#). Indeed, the proposed *a priori* upper bound of 0.6 for κ is only slightly larger than our point estimate of 0.54 for L , the lower bound defined in [Proposition 2.2](#). After accounting for uncertainty over Σ , we find that 26 percent of the posterior density for L lies above 0.6. As such, our framework strongly suggests that the belief $\kappa < 0.6$ is incompatible with the data, and we cannot proceed further under this prior.

We now consider a second restriction that takes 0.6 as a lower bound on κ , while continuing to impose $\rho_{u\xi^*} \in [0, 0.9]$. This restriction places an *upper* bound on the extent of measurement error, ruling out the most extreme possible values of κ . Results for this restriction appear in the third row of [Table 1](#). This restriction does not yield empty identified sets, as we see from the first column of Panel (II). It does however, strongly suggest that settler mortality is an invalid instrument: 70% of the posterior draws for the reduced form parameters exclude $\rho_{u\zeta} = 0$ under the restriction $(\kappa, \rho_{u\xi^*}) \in (0.6, 1] \times [0, 0.9]$. [Figure 2a](#) makes this point in a slightly different way, by depicting the identified set for $(\kappa, \rho_{u\xi^*}, \rho_{u\zeta})$, evaluated at the posterior mean for $\widehat{\Sigma}$, in the region where $\rho_{u\xi^*}$ is positive.¹⁵ The gray region corresponds to $L < \kappa < 0.6$, the largest amount of measurement error consistent with $\widehat{\Sigma}$. We see from the figure that the plane $\rho_{u\zeta} = 0$ only intersects the identified set in the region where measurement error is extremely severe. Moreover, unless $\kappa = L$, $\rho_{u\zeta} = 0$ implies that $\rho_{u\xi^*}$ must be close to zero, in other words that institutions are approximately exogenous. This seems implausible. Indeed, under the restriction $(\kappa, \rho_{u\xi^*}) \in (0.6, 1] \times [0, 0.9]$, depicted in shades of red and blue in [Figure 2a](#), the identified set resides exclusively below the plane $\rho_{u\zeta} = 0$, suggesting that log settler mortality is *negatively* correlated with the unobservables.

[Figure 2a](#) shows that one would need to place high *a priori* probability on implausible regions of the identified set to support the belief that settler mortality is a valid instrument. Because this set is evaluated at a single value of Σ , however, the figure does not account for

¹⁴Based on footnote 19 of the paper, he expressed the belief that at least 40 percent of the measured variation in quality of institutions was likely to be noise.

¹⁵Note that under our Jeffreys prior the posterior mean equals the maximum likelihood estimator.

uncertainty over the reduced form parameters. In contrast, the posterior credible interval for $\rho_{u\zeta}$ in Panel (III) averages both over the posterior for Σ and over the conditional identified sets themselves, via a uniform reference prior.¹⁶ This interval shows that, averaging over reduced form draws, the *relative area* of the conditional identified compatible with a valid instrument is very small. Notice the stark contrast between our credible interval for the parameter $\rho_{u\zeta}$ in Panel (III) and that for the *identified set* for $\rho_{u\zeta}$ in Panel (II). Panel (II) shows that we cannot exclude the possibility that the identified set for $\rho_{u\zeta}$ includes zero, averaged over uncertainty in Σ . In contrast, Panel (III) shows that one would need to place an inordinate amount of *a priori* probability over very small regions of the identified set to support the claim that z is a valid instrument.

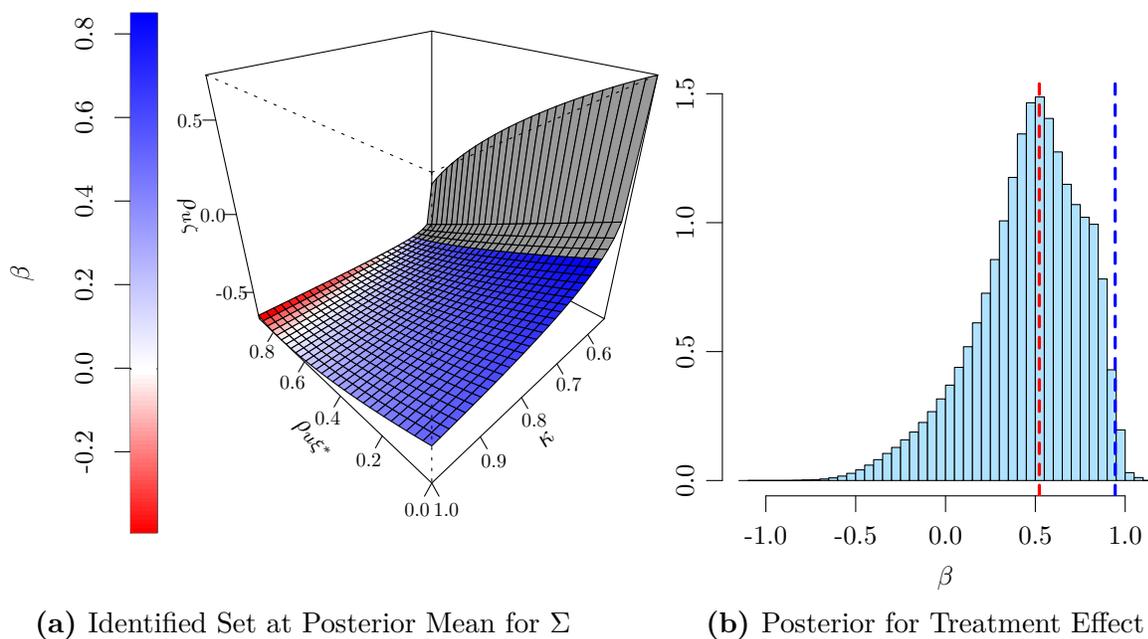


Figure 2: Results for the Colonial Origins example from Section 5.1. Panel (a) plots the identified set for $(\rho_{u\zeta}, \rho_{u\xi^*}, \kappa)$ evaluated at the posterior mean for Σ in the region corresponding to a positive selection effect: $\rho_{u\xi^*} \in [0, 0.9]$. The region in which $0.6 > \kappa$ is shaded in gray while the colors on the remainder of the surface correspond to the implied value of the treatment effect β . Panel (b) gives the posterior for the partially identified parameter β under a uniform prior on the intersection of the restriction $(\kappa, \rho_{u\xi^*}) \in [0.6, 1] \times [0, 0.9]$ with the conditional identified set (see subsection 4.2 for details). The dashed red line gives the OLS estimate and the blue line the IV estimate.

The primary question of interest, of course, is not the validity of settler mortality as an instrumental variable, but the causal effect of institutions on development. The colored region in Figure 2a shows how κ , $\rho_{u\xi^*}$ and $\rho_{u\zeta}$ map into corresponding values for β . Blue indicates a positive treatment effect, red a negative treatment effect, and white a zero treat-

¹⁶See subsection 4.2.

ment effect. In both directions, darker colors indicate larger magnitudes. As seen from the figure, we cannot rule out negative values for β . The posterior credible set for the identified set for β from columns 3–4 of Panel (II) tells the same story, while accounting for sampling uncertainty in Σ . Notice from Figure 2a, however, that at least when evaluated at $\widehat{\Sigma}$, the identified set implies negative values for β only in the region where $\rho_{u\xi^*}$ is extremely large and there is very little measurement error (κ is close to one). Because the posterior for $\underline{\beta}$ is determined *entirely* from these extreme points, the resulting inference is very conservative, a concern that we raised above in subsection 4.2. This observation motivates the idea of averaging not only over reduced form draws Σ but also over the conditional identified set itself, as we do in Panel (III), using a uniform reference prior. Unlike the posterior credible interval for the identified *set* for β in Panel (II), our posterior credible interval for the partially identified *parameter* β , constructed under a conditionally uniform reference prior, contains only positive values.¹⁷ This indicates that the conditional identified sets for $(\kappa, \rho_{u\xi^*}, \rho_{uz})$ contain, on average, only a small region in which β is negative.¹⁸ Indeed, the posterior median for β is 0.49, very close to the OLS estimate from Acemoglu et al. (2001). As we see from 2b, the posterior from which the credible interval in Panel (III) was constructed, the IV estimate is very likely an *overestimate*. In spite of the likely negative correlation between settler mortality and u under reasonable prior beliefs that accord with the data, the main result of Acemoglu et al. (2001) continues to hold: it appears that the effect of institutions on income per capita is almost certainly positive.

5.2 Was Weber Wrong?

We now consider an application in which our framework leads to very different conclusions from those of the preceding example. Becker and Woessmann (2009) study the long-run effect of the adoption of Protestantism in sixteenth-century Prussia on a number of economic and educational outcomes, using variation across counties in their distance to Wittenberg – the city where Martin Luther introduced his ideas and preached – as an instrument for the Protestant share of the population in the 1870s. Here we consider their estimates of the effect of Protestantism on literacy, based on the specification

$$\begin{aligned} \text{Literacy rate} &= \text{constant} + \beta (\text{Protestant share}) + \mathbf{x}'\gamma + u \\ \text{Protestant Share} &= \text{constant} + \pi (\text{Distance to Wittenberg}) + \mathbf{x}'\delta + v \end{aligned}$$

¹⁷See section 4 for a detailed discussion of the difference between inference for the identified set and inference for the partially identified parameter.

¹⁸Because the prior is uniform, “small” refers to the relative area of a region on the identified set: in Figure 2a, for example, the red region is small compared to the blue and white regions.

where \mathbf{x} is a vector of demographic and regional controls.¹⁹

Becker and Woessmann (2009) express beliefs about the three key parameters in our framework. First, their IV strategy relies on the assumption that $\rho_{u\zeta} = 0$, an assumption that we will relax below. Second, the authors argue that the 1870 Prussian Census is regarded by historians to be highly accurate. As such, measurement error in the Protestant share should be fairly small. Finally, Becker and Woessmann (2009) go through a lengthy discussion of the nature of the endogeneity of the Protestant share, suggesting that it is most likely that Protestantism is *negatively* correlated with the unobservables:

wealthy regions may have been less likely to select into Protestantism at the time of the Reformation because they benefited more from the hierarchical Catholic structure, because the opportunities provided by indulgences allured to them, and because the indulgence costs weighted less heavily on them ...The fact that “Protestantism” was initially a “protest” movement involving peasant uprisings that reflected social discontent is suggestive of such a negative selection bias (pp. 556-557).

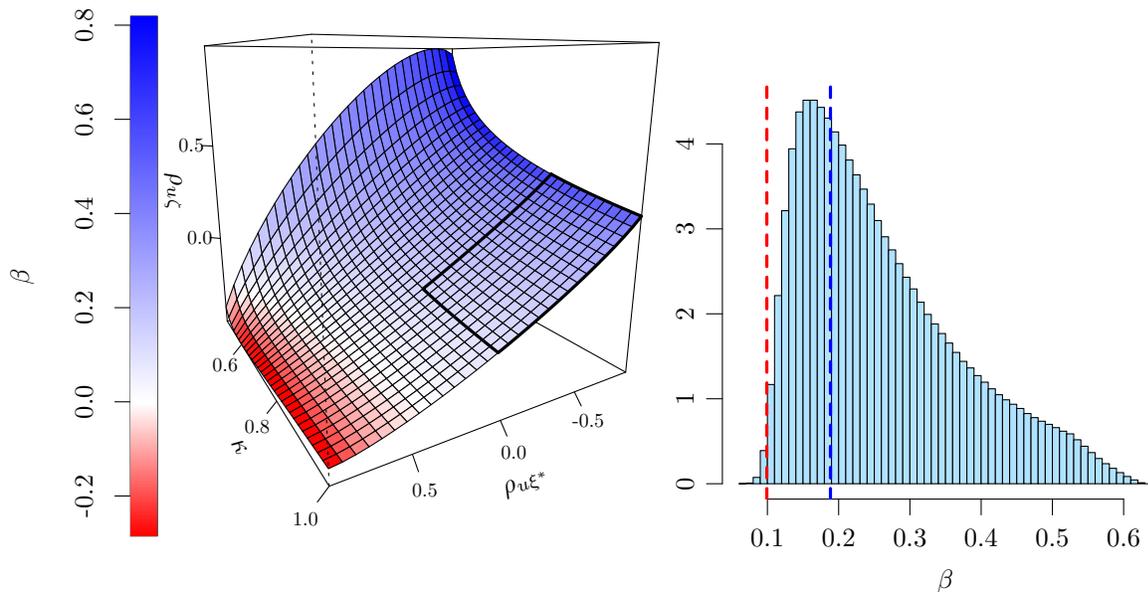
Results for the “Was Weber wrong?” example appear in Table 2. Estimates and bounds for β indicate the percentage point change in literacy that a county would experience if its share of Protestants were to increase by one percentage point. All other values in the table are unitless: they are either probabilities, correlations, or variance ratios. OLS and IV estimates and standard errors, along with the estimates of the lower bounds L for κ appear in row four of Panel (I). Panel (II) presents inference for the identified set. The first column of Panel (II) gives the fraction of posterior draws for the reduced form parameters that yield an empty identified set, while the second column gives the fraction that are compatible with a valid instrument: $\rho_{u\zeta} = 0$. The third and fourth columns of Panel (II) present 90% posterior credible intervals for the identified sets for $\rho_{u\zeta}$ and β , constructed by symmetrically expanding around the conditional identified set evaluated at the posterior mean for Σ , as described in subsection 4.1. In contrast, panel (III) presents posterior medians and 90% highest posterior density intervals for the partially identified parameters $\rho_{u\zeta}$ and β .

As we see from Table 2, Becker and Woessmann (2009) obtain an OLS estimate of 0.10 and an IV estimate that is nearly twice as large: 0.19 with a standard error of 0.03. If the instrument is valid, this corresponds to just under a 0.2 percentage point increase in literacy from each percentage point increase in the prevalence of Protestantism in a given county. The estimated lower bound for κ in this example is just under a half, which means that at most 50

¹⁹In this exercise we include the controls listed in Section III of Becker and Woessmann (2009), specifically: the fraction of the population younger than age 10, of Jews, of females, of individuals born in the municipality, of individuals of Prussian origin, the average household size, log population, population growth in the preceding decade, the fraction of the population with unreported education information, and fraction of the population that was blind, deaf-mute, and insane.

percent of the measured variation in the Protestant share can be attributed to measurement error. Notice that this bound is somewhat weak: it allows for far more measurement error than one might consider reasonable given the author’s arguments concerning the accuracy of the Prussian census data.

Figure 3a depicts the identified set for $(\kappa, \rho_{u\xi^*}, \rho_{u\zeta})$ evaluated at the posterior mean for Σ . As above, the surface is colored to indicate the corresponding value of β : blue indicates a positive treatment effect, red a negative effect, and zero no effect. In both directions, darker colors indicate larger magnitudes. We see immediately from the figure, that unless $\rho_{u\xi^*}$ is large and *positive*, the treatment effect will be positive, irrespective of the amount of measurement error. The rectangular region surrounded by thick black boundaries indicates our approximation to the prior beliefs of Becker and Woessmann (2009): negative selection, and measurement error that is not too severe. This area is well within the blue region, corresponding to a positive treatment effect. Although it is somewhat harder to see from the figure, the region enclosed in the black boundary also contains $\rho_{u\zeta} = 0$. The belief that $\rho_{u\xi^*} < 0$ and measurement error is modest indeed appears to be compatible with a valid instrument in this example.



(a) Identified Set at Posterior Mean for Σ

(b) Posterior for Treatment Effect

Figure 3: Results for the “Was Weber Wrong?” example from Section 5.2. Panel (a) plots the identified set for $(\rho_{u\zeta}, \rho_{u\xi^*}, \kappa)$ evaluated at the posterior mean for Σ . The color of the surface corresponds to the implied value of the treatment effect β . Panel (b) gives the posterior for the partially identified parameter β under a uniform prior on the intersection of the restriction $(\kappa, \rho_{u\xi^*}) \in [0.8, 1] \times [-0.9, 0]$ with the conditional identified set (see subsection 4.2 for details). The dashed red line gives the OLS estimate and the blue line the IV estimate.

Although the substance of this example is apparent from Figure 3a, merely examining the identified set evaluated at the MLE is insufficient, as it fails to account for uncertainty in the reduced form parameters Σ . Row 3 of Table 2 completes our analysis by providing Bayesian inference for both the identified set and the partially identified parameters in the Weber example, imposing the restriction indicated by the black boundary in Figure 3a: $\kappa > 0.8$ and $-0.9 < \rho_{u\xi^*} < 0$. In this example both the inferences for the identified set in Panel (II) and the inferences for the partially identified set in Panel (III) tell the same story: it is extremely unlikely, *a priori*, that β could be negative in this example given the researcher beliefs we have imposed. This is because 100% of the reduced form draws for this prior yield an identified set that contains $\rho_{u\xi} = 0$. Similarly, the posterior median for $\rho_{u\xi}$ under a conditionally uniform reference prior, shown in Panel (III) is very close to zero. If we wish to report a point estimate for β , the posterior median from our uniform reference prior in the second column of Panel (III) suggests that the IV estimate is approximately correct, although the highest posterior density interval is skewed somewhat towards even *larger* causal effects. Moreover, none of these results is sensitive to the restriction $\kappa > 0.8$, as we see from row 2 of Table 2 which imposes only $-0.9 < \rho_{u\xi^*} < 0$. In this example, the authors beliefs are mutually consistent and their result is extremely robust.

5.3 Afghan Girls RCT

Burde and Linden (2013) study the effect of village schools on the academic performance of children in rural northwestern Afghanistan, using data from a randomized controlled trial. Both test scores and reported enrollment rates increased significantly in villages that were randomly allocated to receive a school compared to those that were not. The effects were particularly striking for girls, whose enrollment increased by 52 percentage points and test scores by 0.65 standard deviations. Both effects are statistically significant at the 1 percent level and remain essentially unchanged after controlling for a host of demographic covariates.

These results quantify the causal effect of establishing a school in a rural village. But the data from Burde and Linden (2013) are rich enough for us to pose a more specific question that the authors do not directly address in their paper: what is the causal effect of school attendance on the test scores of Afghan girls? With school enrollment as our treatment of interest, the 0.65 standard deviation increase in test scores becomes an intent to treat (ITT) effect, while the 52 percent increase in reported enrollment becomes an IV first stage. In this example we consider the specification

$$\text{Test score} = \text{constant} + \beta (\text{Enrollment}) + \mathbf{x}'\gamma + \varepsilon$$

and instrument enrollment using the experimental randomization: Girls in a village where a school was established have $z = 1$ and girls in a village where none was have $z = 0$. The vector \mathbf{x} contains the same covariates used by [Burde and Linden \(2013\)](#).²⁰ This dataset has three features that make it an ideal candidate for the methods we have developed above. First, the enrollment variable measures not whether a girl attended the newly-established village school, but whether she attended a school of any kind. This means that our treatment of interest, enrollment, is endogenous: the sample contains 248 girls who did not enroll despite a school being established in their village, and 49 who attended school despite the lack of one in their village. In this example positive selection, $\rho_{u\xi^*} > 0$, seems uncontroversial: parents who enroll their daughter in school are likely to have other unobserved characteristics favorable for their academic performance. Second, although the allocation of village schools was randomized, this does not necessarily make it a valid instrument. Indeed, the authors argue that establishing a village school may affect performance through channels other than increased enrollment alone if, for example,

the village-based schools were of lower quality than the traditional public schools, and some treatment students who would have otherwise attended traditional public schools attended village-based schools instead, or if children who were not enrolled in the treatment group experienced positive spillovers from enrolled siblings or other peers. ([Burde and Linden \(2013\)](#), p. 36.)

Third, school enrollment status is determined from a household survey and, as such, could be subject to substantial mis-reporting. Note that non-differential measurement error in enrollment would not affect the ITT estimate but would bias the estimated causal effect of establishing a school on enrollment.

Results for the Afghan Girls RCT example appear in [Table 3](#). Estimates and bounds for β indicate the standard deviation increase in girls' test scores that would result from enrolling in school. All other values in the table are unitless. The first two columns of Panel (I) present OLS and IV estimates and standard errors. The final three columns of Panel (I) contain posterior means of the upper bounds for the mis-classification probabilities (α_0, α_1) and the *lower* bound for $\psi = -(\alpha_0 + \alpha_1)$. These are calculated by setting $\tilde{\kappa} = L$ and applying [Proposition 3.1](#), and hence correspond to the axis intercepts and point of tangency of the dashed curve in [Figure 1](#). The first column of Panel (II) gives the fraction of posterior draws for the reduced form parameters that yield an empty identified set, while the second

²⁰These are: an indicator for whether the girl is a child of the household head, the girl's age, the number of years the household has lived in the village, a Farsi dummy, a Tajik dummy, a farmers dummy, the age of the household head, years of education of the household head, the number of people in the household, Jeribs of land, number of sheep, distance to the nearest formal school, and a dummy for Chagcharan province.

	(I) Summary Statistics				(II) Inference for Θ			(III) Inference for θ			
	OLS	IV	$\bar{\alpha}_0$	$\bar{\alpha}_1$	ψ	$\mathbb{P}(\emptyset)$	$\mathbb{P}(\text{Valid})$	$\rho_{u\zeta}$	β	$\rho_{u\zeta}$	β
Afghan Girls RCT ($n = 687$)	0.86 (0.06)	1.30 (0.12)	0.24	0.32	-0.30						
$\rho_{u\xi^*} \in [0, 0.9]$						0.00	1.00	[-0.34, 0.65]	[-2.66, 2.47]	0.24	0.51
$\rho_{u\xi^*} \in [0, 0.5]$						0.00	1.00	[-0.33, 0.44]	[-0.13, 2.54]	[-0.15, 0.58]	[-0.92, 1.58]
$\rho_{u\xi^*} \in [0, 0.2]$						0.00	1.00	[-0.33, 0.30]	[0.17, 2.49]	0.11	0.84
$\rho_{u\xi^*} \in [0.5, 0.9]$						0.00	0.00	[0.10, 0.65]	[-2.91, 1.45]	[-0.24, 0.37]	[0.20, 1.49]
										0.01	1.06
										[-0.34, 0.24]	[0.63, 1.57]
										0.44	-0.27
										[0.23, 0.60]	[-1.50, 0.75]

Table 3: Results for the Afghan Girls RCT example. The final three columns of Panel (I) contain posterior means of the upper bounds for the mis-classification probabilities (α_0, α_1) and the *lower* bound for $\psi = -(\alpha_0 + \alpha_1)$. These bounds correspond to the axis intercepts and point of tangency of the dashed curve in [Figure 1](#). Panels (II) and (III) present posterior inferences under interval restrictions on $\rho_{u\xi^*}$. Panel (II) gives posterior inference for the identified set. The column $\mathbb{P}(\emptyset)$ gives the fraction of reduced form parameter draws that yield an empty identified set, while $\mathbb{P}(\text{Valid})$ gives the fraction of reduced form parameter draws compatible with a valid instrument ($\rho_{u\zeta} = 0$). The remaining columns of Panel (II) give 90% posterior credible intervals for the identified set for $\rho_{u\zeta}$ and β . In contrast, Panel (III) presents posterior medians and 90% credible intervals for the partially identified parameters $\rho_{u\zeta}$ and β under a conditionally uniform reference prior. For details, see [section 4](#).

column gives the fraction that are compatible with a valid instrument: $\rho_{u\zeta} = 0$. The third and fourth columns of Panel (II) present 90% posterior credible intervals for the identified sets for $\rho_{u\zeta}$ and β , constructed by symmetrically expanding around the conditional identified set evaluated at the posterior mean for Σ , as described in [subsection 4.1](#). In contrast, panel (III) presents posterior medians and 90% highest posterior density intervals for $\rho_{u\zeta}$ and β , based on the uniform reference prior described in [subsection 4.2](#).

At 0.86 standard deviations, the OLS estimate in this example is quite large, but the IV estimate is even larger: 1.3 standard deviations. The posterior mean for $\underline{\psi}$, the lower bound for $\psi = -(\alpha_0 + \alpha_1)$, however, equals -0.3 . Abstracting from sampling uncertainty in the reduced form parameters, this would imply that $(1 - \alpha_0 - \alpha_1)$ lies in the range $[0.7, 1]$. Hence, if z was a valid instrument, we would obtain a range of approximately $[0.9, 1.3]$ for the true causal effect, via [\(17\)](#): non-differential measurement error in a binary regressor *inflates* the IV estimate. If z is potentially invalid, however, the situation is more complicated. We consider four possible restrictions on regressor endogeneity that impose $\rho_{u\zeta^*} > 0$, corresponding to positive selection into treatment. The first three set $\rho_{u\zeta^*} \in [0, \bar{\rho}]$ for $\bar{\rho} \in \{0.2, 0.5, 0.9\}$, corresponding to a belief about the maximum possible extent of positive selection. As we see from [Table 3](#), we learn very little about $\rho_{u\zeta}$ and β under $\rho_{u\zeta^*} \in [0, 0.9]$, regardless of whether we consider inferences for the identified set Θ in Panel (II), or inferences for the partially identified parameter θ in Panel (III). But $\rho_{u\zeta^*} = 0.9$ would require an extreme degree of positive selection. Lowering the upper bound for $\rho_{u\zeta^*}$ to 0.5 and 0.2, we see that inferences for β become informative. Under $\rho_{u\zeta^*} \in [0, 0.2]$ the 90% posterior credible interval for the identified set for β comfortably excludes zero, as we see from Panel (II). Under both $\rho_{u\zeta^*} \in [0, 0.5]$ and $\rho_{u\zeta^*} \in [0, 0.2]$, the 90% posterior credible interval for β under a conditionally uniform prior suggests a substantial positive return to enrollment. In none of these cases, however, do our inferences for $\rho_{u\zeta}$ indicate whether z is invalid. The last two rows in [Table 3](#) consider an alternative restriction under which $\rho_{u\zeta^*} \in [0.5, 0.9]$. This corresponds to a researcher belief that there is a *very large* degree of positive selection. Under this restriction, the tables are turned: while we can say nothing about β , we have very clear evidence that z is invalid and that $\rho_{u\zeta}$ is positive. Thus, a researcher who believes in a high degree of positive selection would find empirical support for the positive-spillovers story suggested as a possible channel for instrument invalidity in [Burde and Linden \(2013\)](#).

6 Conclusion and Extensions

Causal inference relies on researcher beliefs. The main message of this paper is that imposing them requires a formal framework, both to guard against contradiction and to ensure that

we learn everything that the data have to teach us. While this point is general, we have focused here on a simple but common setting, that of a linear model with a mis-measured, endogenous treatment and a potentially invalid instrument, presenting both results for the case of a continuous treatment subject to classical measurement error and that of a binary treatment subject to non-differential measurement error. By characterizing the relationship between measurement error, treatment endogeneity, and instrument invalidity in terms of intuitive and empirically meaningful parameters, we have developed a Bayesian tool for eliciting, disciplining, and incorporating credible researcher beliefs in the form of sign and interval restrictions. As we have demonstrated through a wide range of illustrative empirical examples, even relatively weak researcher beliefs can be surprisingly informative in practice. The methods we describe above could be extended in a number of directions. One possibility is to allow for multiple instrumental variables. Another would be to allow for heterogeneous treatment effects in a local average treatment effect (LATE) setting.

References

- Acemoglu, D., Johnson, S., Robinson, J. A., 2001. The colonial origins of comparative development: An empirical investigation. *The American Economic Review* 91 (5), 1369–1401.
- Amir-Ahmadi, P., Drautzburg, T., 2019. Identification and inference with ranking restrictions, Working Paper.
- Arias, J. E., Rubio-Ramírez, J. F., Waggoner, D. F., 2018. Inference based on structural autoregressions identified with sign and zero restrictions: Theory and applications.
- Baumeister, C., Hamilton, J. D., September 2015. Sign restrictions, structural vector autoregressions, and useful prior information. *Econometrica* 83 (5), 1963–1999.
- Becker, S. O., Woessmann, L., 2009. Was Weber wrong? A human capital theory of Protestant economic history. *Quarterly Journal of Economics* 124 (2), 531–596.
- Bekker, P., Kapteyn, A., Wansbeek, T., 1987. Consistent sets of estimates for regressions with correlated or uncorrelated measurement errors in arbitrary subsets of all variables. *Econometrica: Journal of the Econometric Society*, 1223–1230.
- Black, D., Berger, M., Scott, F., 2000. Bounding parameter estimates with nonclassical measurement error. *Journal of the American Statistical Association* 95 (451), 739–748.
- Bollinger, C. R., 1996. Bounding mean regressions when a binary regressor is mismeasured. *Journal of Econometrics* 73, 387–399.
- Bollinger, C. R., 2003. Measurement error in human capital and the black-white wage gap. *Review of Economics and Statistics* 85 (3), 578–585.

- Bollinger, C. R., van Hasselt, M., 2017. Bayesian moment-based inference in a regression models with misclassification error. *Journal of Econometrics* 200, 282–294.
- Burde, D., Linden, L., 2013. Bringing education to Afghan girls: A randomized controlled trial of village-based schools. *AEJ: Applied Economics* 5 (3), 27–40.
- Chen, X., Christensen, T., O’Hara, K., Tamer, E., 2016. MCMC confidence sets for identified sets, arXiv:1605.00499.
- Conley, T. G., Hansen, C. B., Rossi, P. E., 2012. Plausibly exogenous. *The Review of Economics and Statistics* 94 (1), 260–272.
- DiTraglia, F., García-Jimeno, C., 2019. Identifying the effect of a mis-classified, binary, endogenous regressor. *Journal of Econometrics* 209 (2), 376–390.
- Frazis, H., Lowenstein, M. A., 2003. Estimating linear regressions with mismeasured, possibly endogenous, binary explanatory variables. *Journal of Econometrics* 117 (1), 151–178.
- Gundersen, C., Kreider, B., Pepper, J., 2012. The impact of the national school lunch program on child health: A nonparametric bounds analysis. *Journal of Econometrics* 166 (1), 79–91.
- Gustafson, P., 2015. Bayesian Inference for Partially Identified Models: Exploring the Limits of Limited Data. No. 141 in *Monographs on Statistics and Applied Probability*. CRC Press, Boca Raton.
- Hahn, P. R., Murray, J. S., Manolopoulou, I., 2016. A Bayesian partial identification approach to inferring the prevalence of accounting misconduct. *Journal of the American Statistical Association* 111 (513).
- Hartigan, J., 1983. *Bayes Theory*. Springer, New York.
- Hu, Y., 2008. Identification and estimation of nonlinear models with misclassification error using instrumental variables: A general solution. *Journal of Econometrics* 144 (1), 27–61.
- Kahneman, D., Tversky, A., 1974. Judgement under uncertainty: Heuristics and biases. *Science* 185 (4157), 1124–1131.
- Kane, T., Rouse, C. E., Staiger, D., July 1999. Estimating the returns to schooling when schooling is misreported, NBER Working Paper # 7235.
- Kitagawa, T., July 2012. Estimation and inference for set-identified parameters using posterior lower probability, Working Paper.
URL <http://www.homepages.ucl.ac.uk/~uctptk0/Research/LowerUpper.pdf>
- Klepper, S., Leamer, E. E., 1984. Consistent sets of estimates for regressions with errors in all variables. *Econometrica* 52 (1), 163–184.
- Kline, B., Tamer, E., July 2016. Bayesian inference in a class of partially identified models. *Quantitative Economics* 7 (2).

- Krasker, W. S., Pratt, J. W., 1986. Bounding the effects of proxy variables on regression coefficients. *Econometrica* 54 (3), 641–655.
- Kreider, B., Pepper, J. V., 2007. Disability and employment: reevaluating the evidence in light of reporting errors. *Journal of the American Statistical Association* 102 (478), 432–441.
- Kreider, B., Pepper, J. V., Gundersen, C., Jolliffe, D., 2012. Identifying the effects of snap (food stamps) on child health outcomes when participation is endogenous and misreported. *Journal of the American Statistical Association* 107 (499), 958–975.
- Leamer, E. E., 1987. Errors of variables in linear systems. *Econometrica* 55 (4), 893–909.
- Lewbel, A., March 2007. Estimation of average treatment effects with misclassification. *Econometrica* 75 (2), 537–551.
- Mahajan, A., 2006. Identification and estimation of regression models with misclassification. *Econometrica* 74 (3), 631–665.
- Moon, H. R., Schorfheide, F., 2009. Estimation with overidentifying inequality moment conditions. *Journal of Econometrics* 153, 136–154.
- Moon, H. R., Schorfheide, F., 2012. Bayesian and frequentist inference in partially identified models. *Econometrica* 80 (2), 755–782.
- Müller, U. K., 2013. Risk of Bayesian inference in misspecified models, and the sandwich covariance matrix. *Econometrica* 81 (5), 1805–1849.
- Nevo, A., Rosen, A. M., 2012. Identification with imperfect instruments. *The Review of Economics and Statistics* 94 (3), 659–671.
- Poirier, D., 1998. Revising beliefs in nonidentified models. *Econometric Theory* 14, 483–509.
- Richardson, T. S., Evans, R. J., Robins, J. A., 2011. Transparent parameterizations of models for potential outcomes. In: *Bayesian Statistics*. Vol. 9. pp. 569–610.
- Sims, C., 2010. Understanding non-Bayesians. Unpublished chapter, Department of Economics, Princeton University.
- van Hasselt, M., Bollinger, C. R., 2012. Binary misclassification and identification in regression models. *Economics Letters* 115, 81–84.
- Zellner, A., 1971. *An Introduction to Bayesian Inference in Econometrics*. John Wiley and Sons, Hoboken, NJ.

Online Appendix

A Framework for Eliciting, Incorporating, and Disciplining Identification Beliefs in Linear Models

Francis J. DiTraglia
University of Oxford

Camilo García-Jimeno
Federal Reserve Bank of Chicago and NBER

A Proofs

Proof of Lemma 2.1. By the definitions of (u, v, ζ, w) and the properties of covariance,

$$\begin{aligned}\sigma_{uw} &= [\text{Cov}(u, \tilde{w}) - \psi \text{Cov}(u, T^*)] \\ \sigma_{\zeta w} &= [\text{Cov}(z, \tilde{w}) - \psi \text{Cov}(z, T^*)] - [\text{Cov}(\tilde{w}, \mathbf{x}') - \psi \text{Cov}(T^*, \mathbf{x}')] \boldsymbol{\varphi}_z \\ \sigma_{vw} &= [\text{Cov}(T^*, \tilde{w}) - \psi \text{Var}(T^*)] - \pi [\text{Cov}(z, \tilde{w}) - \psi \text{Cov}(z, T^*)] - [\text{Cov}(\mathbf{x}', \tilde{w}) - \psi \text{Cov}(\mathbf{x}', T^*)] \boldsymbol{\eta}.\end{aligned}$$

By the definition of ψ , $[\text{Cov}(T^*, \tilde{w}) - \psi \text{Var}(T^*)] = 0$. Moreover, by [Assumption 2.2](#) all of the remaining terms in square brackets likewise equal zero. Thus, $\sigma_{uw} = \sigma_{\zeta w} = \sigma_{vw} = 0$. Next, $\sigma_{v\zeta} = \text{Cov}(v, z) - \text{Cov}(v, \mathbf{x}') \boldsymbol{\varphi}_z = 0$ because $\text{Cov}(v, z)$ and $\text{Cov}(v, \mathbf{x}') = 0$ by [Assumption 2.1](#) (ii). Finally, $\text{Cov}(\mathbf{x}, w) = [\text{Cov}(\mathbf{x}, \tilde{w}) - \psi \text{Cov}(\mathbf{x}, T^*)] = \mathbf{0}$ by [Assumption 2.2](#) and the definition of w . \square

Proof of Lemma 2.2. Substituting (2) and the reduced form for z into (1),

$$y = \mathbf{x}' \boldsymbol{\varphi}_y + \varepsilon = \mathbf{x}' [\beta (\pi \boldsymbol{\varphi}_z + \boldsymbol{\eta}) + \boldsymbol{\gamma}] + [\beta (\pi \zeta + v) + u]$$

by equating with the reduced form equation for y from (4). Similarly, substituting (2) and the reduced form for z into (5) gives

$$T = \mathbf{x}' \boldsymbol{\varphi}_T + \xi = \mathbf{x}' [\tau \mathbf{e}_1 + (1 + \psi)(\pi \boldsymbol{\varphi}_z + \boldsymbol{\eta})] + [(1 + \psi)(\pi \zeta + v) + w]$$

by equating with the reduced form equation for T from (4). Now, $\mathbb{E}(w) = 0$ by construction, and since \mathbf{x} includes a constant, ζ and v are likewise mean zero. The result follows since (ζ, v, u) are uncorrelated with \mathbf{x} by [Assumption 2.1](#) and [Lemma 2.1](#). \square

Proof of Lemma 2.3. The result follows immediately by inspection of (9)–(13) and the

equality

$$\begin{bmatrix} \sigma_u^2 & \sigma'_{uv} & \sigma_{u\zeta} \\ \sigma'_{uv} & (\sigma'_v)^2 & 0 \\ \sigma_{u\zeta} & 0 & \sigma_\zeta^2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \left(\frac{1+\psi}{1+\psi'}\right) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \sigma_u^2 & \sigma_{uv} & \sigma_{u\zeta} \\ \sigma_{uv} & \sigma_v^2 & 0 \\ \sigma_{u\zeta} & 0 & \sigma_\zeta^2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \left(\frac{1+\psi}{1+\psi'}\right) & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

with $(1 + \psi) > 0$ and $(1 + \psi') > 0$. \square

Proof of Proposition 2.1. Substituting (18) into (19) and rearranging, $\tilde{\beta} = (s_{13} - \sigma_{u\zeta})/s_{23}$, while solving (21) for $\tilde{\beta}$ gives $\tilde{\beta} = (s_{12} - \tilde{\sigma}_{u\xi^*})/\tilde{\kappa}s_{22}$. Equating these two expressions,

$$\frac{s_{13} - \sigma_{u\zeta}}{s_{23}} = \frac{s_{12} - \tilde{\sigma}_{u\xi^*}}{\tilde{\kappa}s_{22}}. \quad (\text{A.1})$$

Similarly, substituting $\tilde{\beta} = (s_{13} - \sigma_{u\zeta})/s_{23}$ and $\tilde{\beta}\tilde{\kappa}s_{22} = (s_{12} - \tilde{\sigma}_{u\xi^*})$ into (22),

$$(\sigma_u^2 - s_{11}) + \left(\frac{s_{13} - \sigma_{u\zeta}}{s_{23}}\right)(\tilde{\sigma}_{u\xi^*} + s_{12}) = 0. \quad (\text{A.2})$$

Re-arranging (24) gives $\tilde{\sigma}_{u\xi^*} = \rho_{u\xi^*}\sigma_u(\tilde{\kappa}s_{22})^{1/2}$. Substituting this and $\sigma_{u\zeta} = \sigma_u\rho_{u\zeta}s_{33}$ into (A.1)–(A.2),

$$\frac{s_{13} - \sigma_u\rho_{u\zeta}s_{33}}{s_{23}} = \frac{s_{12} - \rho_{u\xi^*}\sigma_u(\tilde{\kappa}s_{22})^{1/2}}{\tilde{\kappa}s_{22}} \quad (\text{A.3})$$

$$(\sigma_u^2 - s_{11}) + \left(\frac{s_{13} - \sigma_u\rho_{u\zeta}s_{33}}{s_{23}}\right) [\rho_{u\xi^*}\sigma_u(\tilde{\kappa}s_{22})^{1/2} + s_{12}] = 0. \quad (\text{A.4})$$

Substituting (A.3) into (A.4) and re-arranging, we obtain

$$\sigma_u^2 = \frac{s_{11}(\tilde{\kappa} - r_{12}^2)}{\tilde{\kappa}(1 - \rho_{u\xi^*}^2)}. \quad (\text{A.5})$$

The result follows by substituting the positive square root of (A.5) into (A.3) and solving the resulting expression for $\rho_{u\zeta}$. \square

Lemma A.1. Under Assumptions 2.1–2.3,

$$(a) \tilde{\sigma}_v^2 = s_{22}(\tilde{\kappa} - r_{23}^2)$$

$$(b) \rho_{uv} = \tilde{\rho}_{uv} = \frac{\rho_{u\xi^*}\sqrt{\tilde{\kappa}} - \rho_{u\zeta}r_{23}}{\sqrt{\tilde{\kappa} - r_{23}^2}}$$

where r_{23} is as defined in Proposition 2.1, $\rho_{uv} \equiv \text{Cor}(u, v)$, and $\tilde{\rho}_{uv} \equiv \text{Cor}(u, (1 + \psi)v)$.

Proof of Lemma A.1(a). By (18), $r_{23}^2 \equiv \text{Cor}(\xi, \zeta)^2 = \tilde{\pi}^2 s_{33}/s_{22}$. By (11) and (17), $s_{22}\tilde{\kappa} = \tilde{\pi}^2 s_{33} + \tilde{\sigma}_v^2$. The result follows by combining these and re-arranging. \square

Proof of Lemma A.1(b). By (15) and (24),

$$\rho_{u\xi^*} = \left(\frac{\tilde{\sigma}_v}{\sqrt{\tilde{\kappa}s_{22}}} \right) \rho_{uv} + \left(\frac{\tilde{\pi}\sigma_\zeta}{\sqrt{\tilde{\kappa}s_{22}}} \right) \rho_{u\zeta}. \quad (\text{A.6})$$

By manipulating Lemma A.1(a), we obtain $\tilde{\sigma}_v/\sqrt{\tilde{\kappa}s_{22}} = \sqrt{1 - r_{23}^2/\tilde{\kappa}}$. From the proof of Lemma A.1(a), $r_{23}^2 = \tilde{\pi}^2 s_{33}/s_{22}$, so that $\tilde{\pi}\sigma_\zeta/\sqrt{\tilde{\kappa}s_{22}} = r_{23}/\sqrt{\tilde{\kappa}}$. The result follows by substituting these two equalities into (A.6) and solving for ρ_{uv} . Because $\sigma_v^2 > 0$ if and only if $\tilde{\sigma}_v^2 > 0$, and $\tilde{\sigma}_v^2 > 0$ if and only if $\tilde{\kappa} > r_{23}^2$ by Lemma A.1(a), the quantity under the radical is always strictly positive making division by $\sqrt{\tilde{\kappa} - r_{23}^2}$ permissible here. \square

Lemma A.2. Under Assumptions 2.1, 2.2, and 2.3(i), the matrix Ω_{11} defined in Lemma 2.1 is positive definite if and only if $\sigma_u^2, \sigma_v^2, \sigma_\zeta^2 > 0$ and $\rho_{uv}^2 + \rho_{u\zeta}^2 < 1$.

Proof of Lemma A.2. By Lemma 2.1, Ω_{11} is positive definite if and only if

$$\sigma_u^2 > 0 \quad (\text{A.7})$$

$$\sigma_u^2 \sigma_v^2 - \sigma_{uv}^2 > 0 \quad (\text{A.8})$$

$$\sigma_\zeta^2 (\sigma_u^2 \sigma_v^2 - \sigma_{uv}^2) - \sigma_v^2 \sigma_{u\zeta}^2 > 0. \quad (\text{A.9})$$

For the ‘‘if’’ direction, first note that by (A.7) we can rearrange (A.8) to yield $\sigma_v^2 > \sigma_{uv}^2/\sigma_u^2 \geq 0$. Dividing through by σ_v^2 , this implies that $|\rho_{uv}| < 1$. Now, since both σ_u^2 and σ_v^2 are strictly positive, we can divide both sides of (A.9) through by $\sigma_v^2 \sigma_u^2$ to obtain $\sigma_\zeta^2 (1 - \rho_{uv}^2) > \sigma_{u\zeta}^2/\sigma_u^2 \geq 0$. Since $\rho_{uv}^2 < 1$, this implies $\sigma_\zeta^2 > 0$. Thus, dividing (A.9) through by $\sigma_v^2 \sigma_u^2 \sigma_\zeta^2$ and rearranging we find that $\rho_{uv}^2 + \rho_{u\zeta}^2 < 1$. For the ‘‘only if’’ direction, $\rho_{uv}^2 + \rho_{u\zeta}^2 < 1$ implies $\rho_{uv}^2 < 1$. Multiplying both sides by $\sigma_u^2 \sigma_v^2$ gives $\sigma_u^2 \sigma_v^2 \rho_{uv}^2 < \sigma_u^2 \sigma_v^2$ since $\sigma_u^2, \sigma_v^2 > 0$. Substituting $\rho_{uv}^2 = \sigma_{uv}^2/(\sigma_u^2 \sigma_v^2)$ and rearranging implies (A.8). Equation A.9 follows similarly, by multiplying both sides of $\rho_{uv}^2 + \rho_{u\zeta}^2 < 1$ by $\sigma_u^2 \sigma_v^2 \sigma_\zeta^2$ and rearranging. \square

Proof of Proposition 2.2. By Assumption 2.3 (ii), Ω_{11} is positive definite. Thus, by Lemma A.2 $\sigma_v^2, \sigma_u^2, \sigma_\zeta^2 > 0$ and $\rho_{uv}^2 + \rho_{u\zeta}^2 < 1$. Since $\sigma_v^2 > 0$ and $\psi \neq -1$ by Assumption 2.1 (v), it follows that $\tilde{\sigma}_v^2 \equiv (1 + \psi)^2 \sigma_v^2 > 0$. Hence, by Lemma A.1(a), $\tilde{\kappa} > \rho_{T_z}^2$. Similarly, since $\sigma_u^2 > 0$, it follows from Equation A.5 in the proof of Proposition 2.1 that $\tilde{\kappa} > r_{12}^2$. Combining these, we see that $\tilde{\kappa} > \max\{r_{12}^2, r_{23}^2\}$. By Lemma A.1(a), $\rho_{uv}^2 + \rho_{u\zeta}^2 < 1$ is equivalent to

$$\left(\frac{\rho_{u\xi^*} \sqrt{\tilde{\kappa}} - \rho_{u\zeta} r_{23}}{\sqrt{\tilde{\kappa} - r_{23}^2}} \right)^2 + \rho_{u\zeta}^2 < 1 \quad (\text{A.10})$$

Putting the terms of (A.10) over a common denominator and rearranging,

$$\rho_{u\xi^*}^2 + \rho_{u\zeta}^2 - \frac{2\rho_{u\xi^*} \rho_{u\zeta} r_{23}}{\tilde{\kappa}^{1/2}} < \frac{\tilde{\kappa} - r_{23}^2}{\tilde{\kappa}}$$

using the fact that $\tilde{\kappa} > r_{23}^2$. Completing the square,

$$\left(\rho_{u\zeta} - \frac{\rho_{u\xi^*} r_{23}}{\tilde{\kappa}^{1/2}} \right)^2 < (1 - \rho_{u\xi^*}^2) \left(\frac{\tilde{\kappa} - r_{23}^2}{\tilde{\kappa}} \right).$$

Now, using (2.1) to substitute for $(\rho_{u\zeta} - \rho_{u\xi^*}r_{23}/\sqrt{\tilde{\kappa}})$, we find that

$$(r_{12}r_{23} - \tilde{\kappa}r_{13})^2 \left[\frac{1 - \rho_{u\xi^*}^2}{\tilde{\kappa}(\tilde{\kappa} - r_{12}^2)} \right] < (1 - \rho_{u\xi^*}^2) \left(\frac{\tilde{\kappa} - r_{23}^2}{\tilde{\kappa}} \right)$$

Cancelling a factor of $(1 - \rho_{u\xi^*}^2)/\tilde{\kappa}$ from each side and rearranging

$$(r_{12}r_{23} - \tilde{\kappa}r_{13})^2 - (\tilde{\kappa} - r_{12}^2)(\tilde{\kappa} - r_{23}^2) < 0 \quad (\text{A.11})$$

using the fact that $\tilde{\kappa} > r_{12}^2$. Expanding and simplifying,

$$(r_{13}^2 - 1)\tilde{\kappa}^2 + (r_{12}^2 + r_{23}^2 - 2r_{12}r_{23}r_{13})\tilde{\kappa} < 0.$$

Since Σ is positive definite, $r_{13}^2 < 1$. Hence, the preceding inequality defines an interval of values that $\tilde{\kappa}$ *cannot* take on, an interval bounded by the roots of a quadratic function that opens downwards. To determine these roots, we factorize as follows:

$$\tilde{\kappa} \left[(r_{13} - 1)\tilde{\kappa} + (r_{12}^2 + r_{23}^2 - 2r_{12}r_{23}r_{13}) \right] = 0.$$

Thus one root is zero and the other is L . To complete the proof, we show that $L < 1$ and $L > \max\{r_{12}^2, r_{23}^2\}$. For the first claim, note that the positive definiteness of Σ implies

$$1 - r_{12}^2 - r_{23}^2 - r_{13}^2 + 2r_{12}r_{23}r_{13} > 0.$$

Rearranging this inequality using $r_{13}^2 < 1$ establishes $L < 1$. For the second claim notice that (A.11) is violated at $\tilde{\kappa} = \max\{r_{12}^2, r_{23}^2\}$. This combined with the fact that the parabola opens downwards establishes that L is greater than both zero and $\max\{r_{12}^2, r_{23}^2\}$. \square

Proof of Theorem 2.1. Let $(\rho_{u\zeta}, \rho_{u\xi^*}, \tilde{\kappa})$ be any triple satisfying $|\rho_{u\xi^*}| < 1$, $\tilde{\kappa} \in (L, 1]$ and (25). Given this triple, the argument proceeds by constructing errors (u, v, w, ξ^*) and parameter values $(\psi, \tau, \pi, \boldsymbol{\eta}, \boldsymbol{\gamma}, \boldsymbol{\varphi}_T^*, \beta)$ that satisfy Assumptions 2.1–2.3 and generate the observed random variables under (1), (2), and (5). This construction depends on the observable reduced form parameters $(\boldsymbol{\varphi}_y, \boldsymbol{\varphi}_T, \boldsymbol{\varphi}_z)$ and errors $(\varepsilon, \xi, \zeta)$.

The first step constructs w so that $\mathbb{E}(w) = 0$, $\sigma_w^2 = s_{22}(1 - \tilde{\kappa})$, $\text{Cov}(w, \varepsilon) = \text{Cov}(w, \zeta) = 0$, $\text{Cov}(w, \mathbf{x}) = \mathbf{0}$, and $\text{Cov}(w, \xi) = \sigma_w^2$. To this end, let χ be the residual from a projection of ξ on ζ and ε , i.e. $\xi = a\varepsilon + b\zeta + \chi$ with $\text{Cov}(\varepsilon, \chi) = \text{Cov}(\zeta, \chi) = 0$. Next let \mathscr{W} be any random variable with $\mathbb{E}(\mathscr{W}) = 0$ and $\text{Var}(\mathscr{W}) = 1$ that is uncorrelated with χ , ε , ζ , and \mathbf{x} . We define w in terms of χ and \mathscr{W} as

$$w = \left(\frac{1 - \tilde{\kappa}}{1 - L} \right) \chi + \left[\frac{s_{22}(1 - \tilde{\kappa})(\tilde{\kappa} - L)}{1 - L} \right]^{1/2} \mathscr{W}. \quad (\text{A.12})$$

Note that the constants in (A.12) are both well-defined and non-negative, since $L < \tilde{\kappa} \leq 1$ by Proposition 2.2. Now, because \mathbf{x} includes a constant, $(\xi, \zeta, \varepsilon)$ are mean zero and hence $\mathbb{E}(w) = 0$ by construction. Moreover, since χ is by construction uncorrelated with ε and ζ , it follows that $\text{Cov}(w, \varepsilon) = \text{Cov}(w, \zeta) = 0$. Similarly, since χ and \mathscr{W} are both uncorrelated

with \mathbf{x} , so is w . To calculate σ_w^2 and $\text{Cov}(w, \xi)$, note that

$$\text{Var}(\chi) = s_{22} - \begin{bmatrix} s_{12} & s_{23} \end{bmatrix} \begin{bmatrix} s_{11} & s_{13} \\ s_{13} & s_{33} \end{bmatrix}^{-1} \begin{bmatrix} s_{12} \\ s_{23} \end{bmatrix} = s_{22}(1 - L)$$

from which it follows that

$$\sigma_w^2 = \left(\frac{1 - \tilde{\kappa}}{1 - L} \right)^2 s_{22}(1 - L) + \left[\frac{s_{22}(1 - \tilde{\kappa})(\tilde{\kappa} - L)}{1 - L} \right] = s_{22}(1 - \tilde{\kappa})$$

and

$$\text{Cov}(w, \xi) = \text{Cov}(w, a\varepsilon + b\zeta + \chi) = \text{Cov}(w, \chi) = \left(\frac{1 - \tilde{\kappa}}{1 - L} \right) \text{Var}(\chi) = s_{22}(1 - \tilde{\kappa}) = \sigma_w^2.$$

The second step constructs errors (ξ^*, v, u) and parameters $(\varphi_T^*, \boldsymbol{\eta}, \boldsymbol{\gamma})$ so that (1) generates the observed distribution of y , (2) generates a distribution for T^* that is compatible with our observables, and (5) generates the observed distribution of T . To this end, set

$$\xi^* = \frac{\xi - w}{1 + \psi}, \quad v = \frac{\xi - w}{1 + \psi} - \pi\zeta, \quad u = \varepsilon - \beta \left(\frac{\xi - w}{1 + \psi} \right)$$

and

$$\varphi_T^* = \frac{\varphi_T - \tau \mathbf{e}_1}{1 + \psi}, \quad \boldsymbol{\eta} = \frac{\varphi_T - \tau \mathbf{e}_1}{1 + \psi} - \pi \boldsymbol{\varphi}_z, \quad \boldsymbol{\gamma} = \boldsymbol{\varphi}_y - \beta \left(\frac{\varphi_T - \tau \mathbf{e}_1}{1 + \psi} \right).$$

Substituting the preceding expressions along with the reduced forms for T^* and z and simplifying, we obtain

$$\beta T^* + \mathbf{x}' \boldsymbol{\gamma} + u = \mathbf{x}' \boldsymbol{\varphi}_y + \varepsilon, \quad \pi z + \mathbf{x}' \boldsymbol{\eta} + v = \mathbf{x}' \boldsymbol{\varphi}_T^* + \xi^*, \quad \tau + (1 + \psi)T^* + w = \mathbf{x}' \boldsymbol{\varphi}_T + \xi$$

as required. Notice that τ is completely unconstrained in this construction. Moreover, the only restriction imposed on ψ thus far has been $\psi \neq -1$ so that division by $1 + \psi$ is well-defined.

The third step sets π and ψ so that our construction satisfies [Assumption 2.1](#). First, we have

$$\text{Cov}(\mathbf{x}, u) = \text{Cov} \left(\mathbf{x}, \varepsilon - \beta \left[\frac{\xi - w}{1 + \psi} \right] \right) = \mathbf{0}, \quad \text{Cov}(\mathbf{x}, v) = \text{Cov} \left(\mathbf{x}, \frac{\xi - w}{1 + \psi} - \pi\zeta \right) = \mathbf{0}$$

since \mathbf{x} is uncorrelated with the reduced form errors $(\varepsilon, \xi, \zeta)$ by definition, and is likewise uncorrelated with w by construction. This verifies (i) and the first part of (ii). Now set $\pi = s_{23}/[(1 + \psi)s_{33}]$. Since \mathbf{x} is uncorrelated with (ζ, ξ, w) , it follows that

$$\text{Cov}(z, v) = \text{Cov} \left(\mathbf{x}' \boldsymbol{\varphi}_z + \zeta, \frac{\xi - w}{1 + \psi} - \pi\zeta \right) = \frac{s_{23}}{1 + \psi} - \pi s_{33} = 0$$

satisfying the second part of (ii). Since $s_{23} \neq 0$, $\pi \neq 0$ satisfying (iii). Since (iv) simply requires that \mathbf{x} include a constant, this requirement is trivially satisfied. For (v), since

$T = \tau + (1 + \psi)T^* + w$, we have $\text{Cov}(T, T^*) > 0$ for any $\psi > -1$.

The fourth step verifies that our construction satisfies [Assumption 2.2](#). Solving (5) for T^* and combining the result with (3), we obtain $\tilde{w} = \psi(T + \tau + w)/(1 + \psi)$. Accordingly, for any random variable Ξ , we have $\text{Cov}(\Xi, \tilde{w}) = \psi\text{Cov}(\Xi, T + w)/(1 + \psi)$ and $\text{Cov}(\Xi, T^*) = \text{Cov}(\Xi, T - w)/(1 + \psi)$. It follows that $\text{Cov}(\Xi, \tilde{w}) = \psi\text{Cov}(\Xi, T^*)$ if and only if $\text{Cov}(\Xi, w) = 0$. Hence, to verify [Assumption 2.2](#) it suffices to show that $\text{Cov}(u, w) = 0$, $\text{Cov}(z, w) = 0$, and $\text{Cov}(\mathbf{x}, w) = 0$. The first and last of these equalities hold by our construction of w and u above. For the second, we have $\text{Cov}(z, w) = \boldsymbol{\varphi}'_z \text{Cov}(\mathbf{x}, w) + \text{Cov}(\zeta, w) = 0$.

The final step sets $\tilde{\beta} = (s_{13} - \sigma_{u\zeta})/s_{23}$ to ensure that our construction satisfies [Assumption 2.3](#). By [Lemma A.2](#) it suffices to verify that $\sigma_u^2, \sigma_v^2, \sigma_\zeta^2 > 0$ and $\rho_{uv}^2 + \rho_{u\zeta}^2 < 1$. First, $\sigma_\zeta^2 = s_{33} > 0$ since Σ is positive definite. Next,

$$\begin{aligned} \sigma_v^2 &= \text{Var}\left(\frac{\xi - w}{1 + \psi} - \pi\zeta\right) = \left(\frac{1}{1 + \psi}\right)^2 \text{Var}(\xi - w) + \pi^2 s_{33} - 2\left(\frac{\pi}{1 + \psi}\right) \text{Cov}(\xi - w, \zeta) \\ &= \left(\frac{1}{1 + \psi}\right)^2 s_{22}\tilde{\kappa} + \frac{s_{23}^2}{(1 + \psi)^2 s_{33}} - \frac{2s_{23}^2}{(1 + \psi)^2 s_{33}} = \left(\frac{1}{1 + \psi}\right)^2 s_{22}(\tilde{\kappa} - r_{23}^2) \end{aligned}$$

by substituting $\pi = s_{23}/[(1 + \psi)s_{33}]$ and using the properties of w from our construction above. Since $L < \tilde{\kappa} \leq 1$ and $L > r_{23}^2$ by [Proposition 2.2](#), it follows that $\sigma_v^2 > 0$. To establish that $\sigma_u^2 > 0$, we show that our construction satisfies (A.3) and (A.4) from the proof of [Proposition 2.1](#). This implies (A.5) by the argument of [Proposition 2.1](#) and it follows that $\sigma_u^2 > 0$ since $\tilde{\kappa} > r_{12}^2$. To this end, first note that

$$\sigma_u^2 = \text{Var}(\varepsilon) + \left(\frac{\beta}{1 + \psi}\right)^2 \text{Var}(\xi - w) - \frac{2\beta}{1 + \psi} \text{Cov}(\varepsilon, \xi - w) = s_{11} + \tilde{\beta} \left(\tilde{\beta}s_{22}\tilde{\kappa} - 2s_{12}\right) \quad (\text{A.13})$$

To simplify this expression, we use the fact that

$$\sigma_{u\xi^*} \equiv \text{Cov}(u, \xi^*) = \text{Cov}\left(\varepsilon - \beta \left[\frac{\xi - w}{1 + \psi}\right], \frac{\xi - w}{1 + \psi}\right) = \left(\frac{1}{1 + \psi}\right) (s_{12} - \tilde{\beta}s_{22}\tilde{\kappa}).$$

Rearranging, $\tilde{\sigma}_{u\xi^*} \equiv (1 + \psi)\sigma_{u\xi^*} = s_{12} - \tilde{\beta}s_{22}\tilde{\kappa}$. Substituting this into (A.13) along with $\tilde{\beta} = (s_{13} - \sigma_{u\zeta})/s_{23}$ gives (A.2). Solving $\tilde{\sigma}_{u\xi^*} = s_{12} - \tilde{\beta}s_{22}\tilde{\kappa}$ for $\tilde{\beta}$ and equating this with $\tilde{\beta} = (s_{13} - \sigma_{u\zeta})/s_{23}$ gives (A.1). As explained in the proof of [Proposition 2.1](#), (A.3) and (A.4) follow from (A.1) and (A.2) by substituting $\rho_{u\zeta} = \sigma_u \rho_{u\zeta} s_{33}$ and $\tilde{\sigma}_{u\xi^*} = \rho_{u\xi^*} \sigma_u (\tilde{\kappa}s_{22})^{1/2}$. The first of these equalities is simply the definition of $\rho_{u\zeta}$, so it suffices to verify the second. By our construction,

$$\text{Var}(\xi^*) = \text{Var}\left(\frac{\xi - w}{1 + \psi}\right) = \left(\frac{1}{1 + \psi}\right)^2 s_{22}\tilde{\kappa} = \left(\frac{1}{1 + \psi}\right)^2 s_{22} [(1 + \psi)^2 \kappa] = s_{22}\kappa$$

and hence

$$\tilde{\sigma}_{u\xi^*} \equiv (1 + \psi)\sigma_{u\xi^*} = \rho_{u\xi^*} \sigma_u (1 + \psi)\sigma_{\xi^*} = \rho_{u\xi^*} \sigma_u \sqrt{s_{22}(1 + \psi)^2 \kappa} = \rho_{u\xi^*} \sigma_u \sqrt{s_{22}\tilde{\kappa}}$$

as required. All that remains is to verify $\rho_{uv}^2 + \rho_{uz}^2 < 1$. To establish this, we show that our construction satisfies the expression for ρ_{uv} given [Lemma A.1](#) (b). The required inequality then follows, given our choice of $\rho_{u\zeta}$ to satisfy [\(25\)](#), because the steps in the proof of [Proposition 2.2](#) are reversible. By our construction of u and v from above,

$$\sigma_{uv} = \text{Cov} \left(\varepsilon - \beta \left[\frac{\xi - w}{1 + \psi} \right], \left[\frac{\xi - w}{1 + \psi} - \pi \zeta \right] \right) = \left(\frac{1}{1 + \psi} \right) s_{12} - \tilde{\beta} \left(\frac{1}{1 + \psi} \right) s_{22\tilde{\kappa}} - \pi s_{13} + \pi \tilde{\beta} s_{23}.$$

Substituting our choices of π and $\tilde{\beta}$ along with the expression for $\tilde{\sigma}_{u\xi^*}$ used in our derivation of σ_u^2 , this simplifies to

$$\sigma_{uv} = \left(\frac{1}{1 + \psi} \right) \left(\tilde{\sigma}_{u\xi^*} - \frac{s_{23}}{s_{33}} \sigma_{u\zeta} \right).$$

Substituting $\sigma_{uv} = \rho_{uv} \sigma_u \sigma_v$, $\tilde{\sigma}_{u\xi^*} = \rho_{u\xi^*} \sigma_u (s_{22\tilde{\kappa}})^{1/2}$, $\sigma_{u\zeta} = \rho_{u\zeta} \rho_u \sqrt{s_{33}}$ and re-arranging gives

$$\sigma_v \rho_{uv} = \left(\frac{1}{1 + \psi} \right) \left[\rho_{u\xi^*} (s_{22\tilde{\kappa}})^{1/2} - \frac{s_{23}}{\sqrt{s_{33}}} \rho_{u\zeta} \right].$$

The desired result follows since $\sigma_v = [s_{22}(\tilde{\kappa} - r_{23}^2)]^{1/2} / (1 + \psi)$ as shown above. \square

Proof of Corollary 2.1. This argument is a special case of the reasoning from the proof of [Proposition B.1](#) with $\mathcal{R} = (L, 1] \times [-1, 1]$. We rely on one additional fact, namely that $g(L) = -\text{sign}\{r_{12}r_{23} - Lr_{13}\}$ which follows from some simple algebra. First suppose that $r_{12}r_{23} < Lr_{13}$. In this case g is positive for all $x_1 \in (L, 1]$. If x_2^* is interior, then x_1^* is L or 1 . But in this case $g(L) = 1$ so the maximum must occur at (L, x_2^*) . Having found the maximum, we now require the minimum. The minimum could equal $g(1)$. Alternatively it could occur at a corner solution for x_2^* , in which case f simplifies to $f(x_1, 1) = r_{23}/\sqrt{x_1}$ or $f(x_1, -1) = -r_{23}/\sqrt{x_1}$ depending on whether x_2 equals 1 or -1 . One of these two functions is negative. In contrast, $g(1)$ is positive so it cannot be the minimum: by inspection the minimum occurs at $-|r_{23}|/\sqrt{L}$. Analogous reasoning holds in the case of $r_{12}r_{23} > Lr_{13}$. If $r_{21}r_{23} = Lr_{13}$, then $f(x_1, x_2) = x_2 r_{23} / \sqrt{x_1}$ so we can again find the extrema by inspection. \square

Proof of Corollary 2.2. See the proof of [Proposition B.2](#), of which [Corollary 2.2](#) is a special case. \square

Proof of Lemma 3.1. By the law of total probability,

$$\begin{aligned} \text{Cov}(T^*, T) &= (1 - \alpha_1)p^* - pp^* = \{(1 - \alpha_1) - [\alpha_0(1 - p^*) + (1 - \alpha_1)p^*]\} p^* \\ &= p^*(1 - p^*)(1 - \alpha_0 - \alpha_1) = \text{Var}(T^*)(1 - \alpha_0 - \alpha_1) \end{aligned}$$

and therefore

$$\psi = \frac{\text{Cov}(T^*, \tilde{w})}{\text{Var}(T^*)} = \frac{\text{Cov}(T^*, T)}{\text{Var}(T^*)} - 1 = \frac{\text{Var}(T^*)(1 - \alpha_0 - \alpha_1)}{\text{Var}(T^*)} - 1 = -(\alpha_0 + \alpha_1)$$

by the definition of \tilde{w} from [\(3\)](#), establishing part (i). For part (ii), first note that \tilde{w} can only

take on the values $\{-1, 0, 1\}$ yielding

$$\begin{aligned}\mathbb{E}[\tilde{w}] &= \mathbb{P}(\tilde{w} = 1) - \mathbb{P}(\tilde{w} = -1) = \mathbb{P}(T = 1, T^* = 0) - \mathbb{P}(T = 0, T^* = 1) \\ &= \alpha_0(1 - p^*) - \alpha_1 p^* = \alpha_0 - (\alpha_0 + \alpha_1)p^*\end{aligned}$$

from which we obtain

$$\tau \equiv \mathbb{E}[\tilde{w}] - \psi \mathbb{E}[T^*] = [\alpha_0 - (\alpha_0 + \alpha_1)p^*] + (\alpha_0 + \alpha_1)p^* = \alpha_0.$$

Finally,

$$w \equiv \tilde{w} - \tau - \psi T^* = (T - T^*) - \alpha_0 + (\alpha_0 + \alpha_1)T^* = (T - \alpha_0) - (1 - \alpha_0 - \alpha_1)T^*$$

establishing (iii). \square

Proof of Lemma 3.2. By the law of total probability, $p = \alpha_0(1 - p^*) + (1 - \alpha_1)p^*$. Rearranging this equality gives (i). For part (ii), first notice that $\sigma_w^2 = \mathbb{E}(w^2)$ since w is mean zero by construction. Now, using Lemma 3.1 (iii) the probability mass function of w is

$$\begin{aligned}\mathbb{P}(w = -\alpha_0) &= \mathbb{P}(T = 0, T^* = 0) = (1 - \alpha_0)(1 - p^*) \\ \mathbb{P}(w = \alpha_1 - 1) &= \mathbb{P}(T = 0, T^* = 1) = \alpha_1 p^* \\ \mathbb{P}(w = 1 - \alpha_0) &= \mathbb{P}(T = 1, T^* = 0) = \alpha_0(1 - p^*) \\ \mathbb{P}(w = \alpha_1) &= \mathbb{P}(T = 1, T^* = 1) = (1 - \alpha_1)p^*\end{aligned}$$

and accordingly we have

$$\begin{aligned}\mathbb{E}(w^2) &= \alpha_0^2(1 - \alpha_0)(1 - p^*) + (1 - \alpha_1)^2 \alpha_1 p^* + (1 - \alpha_0)^2 \alpha_0(1 - p^*) + \alpha_1^2(1 - \alpha_1)p^* \\ &= p^* \alpha_1(1 - \alpha_1) + (1 - p^*) \alpha_0(1 - \alpha_0)\end{aligned}$$

after expanding and simplifying. Eliminating p^* using part (i) gives

$$\sigma_w^2 = \frac{1}{1 - \alpha_0 - \alpha_1} [(p - \alpha_0)\alpha_1(1 - \alpha_1) + (1 - p - \alpha_1)\alpha_0(1 - \alpha_0)]$$

from which (ii) follows after straightforward but tedious algebra. \square

Proof of Proposition 3.1. To begin we show that p^* cannot equal zero or one. By Assumption 2.1 (iv), ξ^* must be identically zero if $p^* \in \{0, 1\}$. But since $\xi^* = \pi\zeta + v$ by Equation 14, this can only occur if $|\text{Cor}(\zeta, v)| = 1$ which is ruled out by Assumption 2.3. Similarly, the positive definiteness of Σ implies that $p \notin \{0, 1\}$. Now, solving Lemma 3.2 (b) for α_0 and α_1 in turn, we obtain

$$\alpha_0 = \frac{\sigma_w^2 - p\alpha_1}{1 - p - \alpha_1}, \quad \alpha_1 = \frac{\sigma_w^2 - (1 - p)\alpha_0}{p - \alpha_0}.$$

where $\sigma_w^2 = s_{22}(1 - \tilde{\kappa})$ by (20). By Lemma 3.2 (a) it follows that $\alpha_0 < p$ and $\alpha_1 < 1 - p$

since $0 < p^* < 1$, so neither denominator can be zero. Now, viewing α_1 as a function of α_0 ,

$$\frac{\partial}{\partial \alpha_0} \alpha_1 = \frac{\sigma_w^2 - p(1-p)}{(p - \alpha_0)^2}, \quad \frac{\partial^2}{\partial \alpha_0^2} \alpha_1 = 2 \left[\frac{\sigma_w^2 - p(1-p)}{(p - \alpha_0)^3} \right]$$

so we see that the signs of the first and second derivatives are entirely determined by the sign of $\sigma_w^2 - p(1-p)$. Since $T = \tau + (1 + \psi)T^* + w$ where $\text{Cov}(T^*, w) = 0$, it follows that

$$\text{Var}(T) = p(1-p) = (1 + \psi)^2 \text{Var}(T^*) + \text{Var}(w) = (1 - \alpha_0 - \alpha_1)^2 p^*(1 - p^*) + \sigma_w^2$$

Since $p^* \notin \{0, 1\}$, we have $\sigma_w^2 - p(1-p) < 0$. Thus α_1 is a strictly decreasing and strictly concave function of α_0 on the interval $\alpha_0 \in [0, p)$. Evaluating this function at $\alpha_0 = 0$ we obtain $\alpha_1 = s_{22}(1 - \tilde{\kappa})/p$. Setting $\alpha_1 = 0$ and solving for α_0 , we obtain $\alpha_0 = s_{22}(1 - \tilde{\kappa})/(1-p)$. These are the α_0 and α_1 axis intercepts, respectively, in [Figure 1](#). Note that both are non-negative since $s_{22} \geq 0$ and $\tilde{\kappa} \leq 1$. Since s_{22} is the variance of the residual from a projection of T on \mathbf{x} , we know that $s_{22} \leq p(1-p)$. And because $0 \leq L \leq 1$, it follows that $s_{22}(1-L)/(1-p) \leq p$ and similarly that $s_{22}(1-L)/p \leq (1-p)$. Hence,

$$0 \leq \alpha_0 < s_{22}(1-L)/(1-p) < p, \quad 0 \leq \alpha_1 < s_{22}(1-L)/p < 1-p.$$

These two values cannot occur simultaneously, however. For any value of σ_w^2 less than $s_{22}(1-L)$ the curve relating α_0 and α_1 necessarily lies below the function $E(\alpha_0) = [s_{22}(1-L) - (1-p)]\alpha_0/(p - \alpha_0)$, since

$$\alpha_1 = \frac{\sigma_w^2 - (1-p)\alpha_0}{p - \alpha_0} < \frac{s_{22}(1-L) - (1-p)\alpha_0}{p - \alpha_0}.$$

The function E is the outer envelope given by the dashed black curve in [Figure 1](#), which cannot actually be attained since $\tilde{\kappa} > L$ by [Proposition 2.2](#). Fixing $\tilde{\kappa}$ determines a functional relationship between α_0 and α_1 . To find the corresponding bound for ψ we use the fact that $\psi = -(\alpha_0 + \alpha_1)$ by [Lemma 3.1](#) (i). Since α_1 is a strictly concave function of α_0 , the minimum value of $\alpha_0 + \alpha_1$ is a corner solution: either $s_{22}(1 - \tilde{\kappa})/p$ or $s_{22}(1 - \tilde{\kappa})/(1-p)$ depending on whether p is larger than $1-p$. Again because the function is strictly concave, the maximum value of $\alpha_0 + \alpha_1$ could be either interior or occur at the *opposite* corner. To solve for an interior maximum, we substitute the constraint $\alpha_1 = [s_{22}(1 - \tilde{\kappa}) - (1-p)\alpha_0]/(p - \alpha_0)$ into the objective function to yield

$$(\alpha_0 + \alpha_1) = \alpha_0 + [s_{22}(1 - \tilde{\kappa}) - (1-p)\alpha_0]/(p - \alpha_0)$$

Differentiating the right hand side with respect to α_0 gives the first order condition

$$(p - \alpha_0)^2 + s_{22}(1 - \tilde{\kappa}) - p(1-p) = 0.$$

which is a quadratic in α_0 with roots $\alpha_0 = p \pm \sqrt{p(1-p) - s_{22}(1 - \tilde{\kappa})}$. Since $p(1-p) > \sigma_w^2$ both of these are real. However, the “+” root violates the constraint $\alpha_0 < p$, hence the unique solution is the “-” root. Substituting this into the constraint, we obtain the corresponding

solution for α_1 . Hence, an interior maximum of $(\alpha_0 + \alpha_1)$ occurs at

$$\alpha_0 = p - \sqrt{p(1-p) - s_{22}(1-\tilde{\kappa})}, \quad \alpha_1 = (1-p) - \sqrt{p(1-p) - s_{22}(1-\tilde{\kappa})}$$

Note that $\alpha_0 \geq 0$ iff $s_{22}(1-\tilde{\kappa}) > p(1-2p)$. Similarly, $\alpha_1 \geq 0$ iff $s_{22}(1-\tilde{\kappa}) > (1-p)(2p-1)$. Hence, the maximum value of $(\alpha_0 + \alpha_1)$ is interior iff $s_{22}(1-\tilde{\kappa}) > m(p)$, in which case $(\alpha_0 + \alpha_1) = 2\sqrt{p(1-p)s_{22}(1-\tilde{\kappa})} - 1$. \square

B Additional Results

This appendix contains proofs of two additional results that do not appear in our paper but are used in the associated R package `ivdoctr`: **Proposition B.1** and **Proposition B.2**. These propositions provide bounds for $\rho_{u\zeta}$ and β under interval restrictions on $\tilde{\kappa}$ and $\rho_{u\xi^*}$.

Proposition B.1. *Suppose that $(\tilde{\kappa}, \rho_{u\xi^*})$ is known, a priori, to lie in a set \mathcal{R} that takes the form $\mathcal{R} \equiv [\tilde{a}, \tilde{b}] \times [c^*, d^*] \subseteq (L, 1] \times [-1, 1]$. Then, under the conditions of **Theorem 2.1**, the sharp identified set for $\rho_{u\zeta}$ is the open interval $(\min_S f, \max_S f)$ where*

$$f(\tilde{\kappa}, \rho_{u\xi^*}) \equiv \frac{r_{23}\rho_{u\xi^*}}{\tilde{\kappa}^{1/2}} - (r_{12}r_{23} - r_{13}\tilde{\kappa}) \left[\frac{1 - \rho_{u\xi^*}^2}{\tilde{\kappa}(\tilde{\kappa} - r_{12}^2)} \right]^{1/2}$$

and S is a finite set defined by $S = S_1 \cup S_2 \cup \left\{ [\tilde{a}, \tilde{b}] \times \{c^*, d^*\} \right\}$ where S_1 is given by

$$S_1 \equiv \mathcal{R} \cap \left\{ (\tilde{a}, h(\tilde{a})), (\tilde{b}, h(\tilde{b})) \right\}, \quad h(\tilde{\kappa}) \equiv \frac{-r_{23}(\tilde{\kappa} - r_{12}^2)^{1/2}}{[(r_{12}r_{23} - \tilde{\kappa}r_{13})^2 + r_{23}^2(\tilde{\kappa} - r_{12}^2)]^{1/2}}$$

and S_2 is given by

$$S_2 \equiv \mathcal{R} \cap \left(\{\Xi(c^*) \times \{c^*\}\} \cup \{\Xi(d^*) \times \{d^*\}\} \right)$$

where $\Xi(c^*)$ and $\Xi(d^*)$ denote the set of roots of

$$(1 - \rho_{u\xi^*}^2) [(2r_{12}r_{23} - r_{13}r_{12}^2)\tilde{\kappa} - r_{23}r_{12}^3]^2 - \rho_{u\xi^*}^2 r_{23}(\tilde{\kappa} - r_{12})^3 = 0$$

with $\rho_{u\xi^*}$ held fixed at c^* and d^* , respectively.

Proof of Proposition B.1. To simplify the notation in this argument, we adopt the shorthand $x_1 \equiv \tilde{\kappa}$ and $x_2 \equiv \rho_{u\xi^*}$ and accordingly write $f(x_1, x_2)$ in place of $f(\tilde{\kappa}, \rho_{u\xi^*})$. Similarly, we write $[a, b]$ and $[c, d]$ in place of $[\tilde{a}, \tilde{b}]$ and $[c^*, d^*]$. Let (x_1^*, x_2^*) be an extremum of f and define $\hat{x}_1 = r_{12}r_{23}/r_{13}$. There are two possibilities: either x_2^* is interior or it lies on the boundary. We begin by showing that if x_2^* is interior, x_1^* must lie on the boundary.

If x_2^* is interior, then it must satisfy the first order condition

$$\frac{\partial f(x_1, x_2)}{\partial x_2} = \frac{r_{23}}{\sqrt{x_1}} + \left[\frac{(r_{12}r_{23} - r_{13}x_1)}{\sqrt{x_1}(x_1 - r_{12}^2)} \right] \left[\frac{x_2}{\sqrt{1 - (x_2)^2}} \right] = 0.$$

We can assume $x_1^* \neq \hat{x}_1$, since $x_1^* = \hat{x}_1$ implies $f(x_1, x_2) = r_{23}x_2/\hat{x}_1^{1/2}$, in which case x_2^* could not be interior. Solving the first-order condition, we obtain $x_2^* = h(x_1^*)$ where

$$h(x_1) = \frac{-r_{23}(x_1 - r_{12}^2)^{1/2}}{[(r_{12}r_{23} - r_{13}x_1)^2 + r_{23}^2(x_1 - r_{12}^2)]^{1/2}},$$

eliminating an extraneous solution by noting that x_2^* must have the same sign as the ratio $-r_{23}/(r_{12}r_{23} - r_{13}x_1^*)$. Concentrating x_2 out of f , we have

$$g(x_1) \equiv f(x_1, h(x_1)) = -\text{sign}\{r_{12}r_{23} - x_1r_{13}\} \sqrt{\frac{(r_{12}r_{23} - x_1r_{13})^2 + r_{23}(x_1 - r_{12}^2)}{x_1(x_1 - r_{12}^2)}}.$$

Differentiating and simplifying gives

$$g'(x_1) = -\frac{(L - r_{12}^2)(1 - r_{13}^2)}{2g(x_1)(x_1 - r_{12}^2)}$$

There are three cases. If $b < \hat{x}_1$, then g' is strictly positive on $[a, b]$ and hence g is monotonically increasing on this interval, implying that x_1^* must lie on the boundary. If instead $\hat{x}_1 < a$, then g' is strictly negative on $[a, b]$ and hence g is monotonically decreasing on this interval, likewise implying that x_1^* must lie on the boundary. The remaining case is $a \leq \hat{x}_1 \leq b$. Note that g is strictly increasing for $x_1 \in [a, \hat{x}_1)$ and strictly decreasing for $x_1 \in (\hat{x}_1, b]$. In this case we obtain candidate minima at $x_1 = a$ and $x_1 = b$ but not candidate maxima. This completes our characterization of candidate extrema for interior x_2^* .

Now suppose that x_2^* occurs at a corner. One possibility is that x_1^* likewise occurs at a corner; the other is that x_1^* is interior. In the latter case, it must satisfy the first order condition

$$\frac{\partial f(x_1, x_2)}{\partial x_1} = \frac{-r_{23}x_2}{2x_1^{2/3}} + \left\{ \frac{r_{13}}{\sqrt{x_1(x_1 - r_{12}^2)}} + \frac{(r_{12}r_{23} - x_1r_{13})(2x_1 - r_{12}^2)}{2[x_1(x_1 - r_{12}^2)]^{3/2}} \right\} \sqrt{1 - x_2^2}$$

and hence the roots of the polynomial

$$(1 - x_2^2)[(2r_{12}r_{23} - r_{13}r_{12}^2)x_1 - r_{23}r_{12}^3]^2 - x_2^2r_{23}^2(x_1 - r_{12}^2)^3 = 0$$

holding x_2 fixed at c and d are likewise candidate extrema. Finally, since f is a continuous function, any value of $\rho_{u\zeta}$ within the resulting bounds can be attained. \square

Proposition B.2. *Suppose that $(\tilde{\kappa}, \rho_{u\zeta^*})$ is known, a priori, to lie in a set \mathcal{R} that takes the form $\mathcal{R} \equiv [\tilde{a}, \tilde{b}] \times [c^*, d^*] \subseteq (L, 1] \times [-1, 1]$. Then, under the conditions of [Theorem 2.1](#), the sharp identified set for β is given by*

$$\mathcal{B} = \begin{cases} (-\infty, \infty), & \text{if } [c^*, d^*] = [-1, 1] \\ \left[\frac{s_{13}}{s_{23}} - \max_Q g, \frac{s_{13}}{s_{23}} - \min_Q g \right], & \text{otherwise} \end{cases}$$

where

$$g(\tilde{\kappa}, \rho_{u\xi^*}) \equiv \frac{\sqrt{s_{11}s_{33}}}{\tilde{\kappa}s_{23}} \left[r_{23} \sqrt{\tilde{\kappa} - r_{12}^2} \left(\frac{\rho_{u\xi^*}}{\sqrt{1 - \rho_{u\xi^*}^2}} \right) - (r_{12}r_{23} - \tilde{\kappa}r_{13}) \right]$$

and Q is a finite set defined by $Q = Q_1 \cup \{ \{\tilde{a}, \tilde{b}\} \times \{c^*, d^*\} \}$ where Q_1 is given by

$$Q_1 \equiv \mathcal{R} \cap (\{\Psi(c^*) \times \{c^*\}\} \cup \{\Psi(d^*) \times \{d^*\}\})$$

with

$$\Psi(\rho_{u\xi^*}) = \left\{ 2r_{12}^2 \left(1 - \sqrt{1 - \rho_{u\xi^*}^2} \right) / \rho_{u\xi^*}^2, 2r_{12}^2 \left(1 + \sqrt{1 - \rho_{u\xi^*}^2} \right) / \rho_{u\xi^*}^2 \right\}.$$

Proof of Proposition B.2. To simplify the notation in this argument, we adopt the shorthand $x_1 \equiv \tilde{\kappa}$ and $x_2 \equiv \rho_{u\xi^*}$ and accordingly write $g(x_1, x_2)$ in place of $g(\tilde{\kappa}, \rho_{u\xi^*})$. Similarly, we write $[a, b]$ and $[c, d]$ in place of $[\tilde{a}, \tilde{b}]$ and $[c^*, d^*]$.

Begin by noticing that for any fixed x_1 , g is a strictly monotonic function of x_2 . This means that the extrema of g lie on the boundary for x_2 . Suppose first that $[c, d] = [-1, 1]$. If $r_{23} > 0$, g is strictly increasing in x_2 and for any $x_1 \in (L, 1]$ we have $\lim_{x_2 \rightarrow -1} = -\infty$ and $\lim_{x_2 \rightarrow 1} = +\infty$. If $r_{23} < 0$, then g is strictly decreasing and the limits are reversed. Hence $\mathcal{B} = (-\infty, \infty)$. Now suppose that $[c, d]$ is a strict subset of $(-1, 1)$. In this case we characterize the optimal values of x_1 at $x_2 = c$ and $x_2 = d$. Since $g(x_1, 0) = \sqrt{s_{11}}(r_{13} - r_{12}r_{23}/x_1)$, the extrema of g as a function of x_1 when $x_2 = 0$ occur at a and b . If instead $x_2 \neq 0$, the extrema could still occur at a and b , or they could be interior. If interior, they must satisfy the first order condition

$$x_1^2/4 - r_{12}^2 x_1/x_2^2 + r_{12}^4/x_2^2 = 0$$

yielding the set of solutions

$$\Psi(x_2) = \left\{ 2r_{12}^2 \left(1 - \sqrt{1 - x_2^2} \right) / x_2^2, 2r_{12}^2 \left(1 + \sqrt{1 - x_2^2} \right) / x_2^2 \right\}.$$

Hence, it suffices to evaluate g at all elements of $\mathcal{R} \cap (\{\Psi(c) \times \{c\}\} \cup \{\Psi(d) \times \{d\}\})$ and at the corners $\{a, b\} \times \{c, d\}$. Since g is a continuous function, any point within the bounds for β can be attained. \square

Proposition B.3. Suppose that $(\tilde{\kappa}, \rho_{u\xi^*})$ is known, a priori, to lie in $\mathcal{R} \equiv [\tilde{a}, \tilde{b}] \times [c^*, d^*] \subseteq (L, 1] \times [-1, 1]$. Then, under the conditions of [Proposition 3.1](#),

$$\min_{[\tilde{\alpha}, \tilde{b}]} \underline{\beta}(\tilde{\kappa}) \leq \beta \leq \max_{[\tilde{\alpha}, \tilde{b}]} \overline{\beta}(\tilde{\kappa})$$

where $\underline{\beta}(\tilde{\kappa}) \equiv \min B(\tilde{\kappa})$, $\overline{\beta}(\tilde{\kappa}) \equiv \max B(\tilde{\kappa})$,

$$B(\tilde{\kappa}) = \left\{ (1 + \psi)(s_{13}/s_{23} - g) : \psi \in \{ \underline{\psi}(\tilde{\kappa}), \overline{\psi}(\tilde{\kappa}) \}, g \in \{ \underline{g}(\tilde{\kappa}), \overline{g}(\tilde{\kappa}) \} \right\}$$

$$\underline{g}(\tilde{\kappa}) = \min \{ g(\tilde{\kappa}, c^*), g(\tilde{\kappa}, d^*) \}$$

$$\overline{g}(\tilde{\kappa}) = \max \{ g(\tilde{\kappa}, c^*), g(\tilde{\kappa}, d^*) \}$$

and g is as defined in [Proposition B.2](#)

Proof of Proposition B.3. This follows from [Proposition 3.1](#) along with the fact that g is monotonic in $\rho_{u\xi^*}$ for fixed $\tilde{\kappa}$ and $\beta = (1 + \psi)[s_{13}/s_{23} - g(\tilde{\kappa}, \rho_{u\xi^*})]$. \square

C Uniform Draws on the Conditional Identified Set

In this appendix we provide details of our method for making uniform draws on $\Theta(\varphi^{(j)})$, an ingredient of our procedure for carrying out inference for θ from [subsection 4.2](#). We first describe the classical measurement error case and then explain what changes in the case of a binary T^* . In the classical measurement error case, $\psi = 0$ so that $\tilde{\kappa} = \kappa$. Thus, equation [\(25\)](#) describes a manifold relating $\rho_{u\zeta}$, $\rho_{u\xi^*}$ and κ . To draw uniformly on this manifold, subject to researcher beliefs, we proceed as follows. Let \mathcal{R} denote a rectangular region encoding interval restrictions on κ and $\rho_{u\xi^*}$. We first draw uniformly on \mathcal{R} , and then re-weight these draws based on the local surface area of the manifold at each draw $(\rho_{u\xi^*}^{(\ell)}, \kappa^{(\ell)})$. By local surface area we refer to

$$M(\rho_{u\xi^*}, \kappa) = \sqrt{1 + \left(\frac{\partial \rho_{u\xi}}{\partial \rho_{u\xi^*}}\right)^2 + \left(\frac{\partial \rho_{u\zeta}}{\partial \kappa}\right)^2}. \quad (\text{C.1})$$

The derivatives required to evaluate the function M are

$$\begin{aligned} \frac{\partial \rho_{u\zeta}}{\partial \rho_{u\xi^*}} &= \frac{\rho_{Tz}}{\sqrt{\kappa}} + \frac{\rho_{u\xi^*}(r_{12}r_{23} - \kappa r_{13})}{\sqrt{\kappa(\kappa - r_{12}^2)(1 - \rho_{u\xi^*}^2)}} \\ \frac{\partial \rho_{u\zeta}}{\partial \kappa} &= -\frac{\rho_{u\xi^*}r_{23}}{2\kappa^{3/2}} + \sqrt{\frac{1 - \rho_{u\xi^*}^2}{\kappa(\kappa - r_{12}^2)}} \left\{ r_{13} + \frac{1}{2}(r_{12}r_{23} - \kappa r_{13}) \left[\frac{1}{\kappa} + \frac{1}{\kappa - r_{12}^2} \right] \right\}. \end{aligned}$$

To accomplish the re-weighting, we first evaluate $M^{(\ell)} = M(\rho_{u\xi^*}^{(\ell)}, \kappa^{(\ell)})$ at each draw ℓ that was accepted in the first step. We then calculate $M_{max} = \max_{\ell=1, \dots, L} M^{(\ell)}$ and *resample* the draws $(\rho_{u\zeta}^{(\ell)}, \rho_{u\xi^*}^{(\ell)}, \kappa^{(\ell)})$ with probability $p^{(\ell)} = M^{(\ell)}/M_{max}$. Now suppose that T^* is binary, so that the measurement error is not classical. In this case we proceed in two steps. First, we generate draws on the manifold relating $(\rho_{u\xi^*}, \rho_{u\zeta}, \tilde{\kappa})$ *exactly* as in the classical measurement error case, by simply replacing κ with $\tilde{\kappa}$ in the preceding equations. Given a draw $(\rho_{u\zeta}^{(\ell)}, \rho_{u\xi^*}^{(\ell)}, \tilde{\kappa}^{(\ell)})$ we then generate the corresponding $\psi^{(\ell)}$ by drawing uniformly on the interval $[\underline{\psi}(\tilde{\kappa}^{(\ell)}), \overline{\psi}(\tilde{\kappa}^{(\ell)})]$ defined in [Proposition 3.1](#).

D Bayesian versus Frequentist Inference

Under certain assumptions our inferences for the identified set from [subsection 4.1](#) can be given a Frequentist repeated-sampling interpretation in the limit under the posterior for Σ described in [subsection 4.4](#). We now give a brief overview of how this can be achieved, appealing to results from [Kline and Tamer \(2016\)](#). Alternatively, one could follow the closely related approach of [Kitagawa \(2012\)](#).

Let φ_0 denote the “true” value of the reduced form parameter vector, i.e. the solution to the population maximum likelihood criterion function. In our example, this corresponds to the true reduced form covariance matrix Σ . Under weak regularity conditions on the true data generating process for (y, T, \mathbf{x}, z) , our inverse-Wishart posterior is consistent for φ_0 by Doob’s Theorem.²¹ Now let $\hat{\varphi}_n$ denote the maximum likelihood estimator based on a sample of n observations. In our example this corresponds to the sample covariance matrix $S/(n - k)$ of the regression residuals $Y - X\hat{B}$. Because our prior is continuous with full support and our posterior is consistent for φ_0 , [Hartigan \(1983\)](#) Theorem 11.2 establishes that $\sqrt{n}(\varphi - \hat{\varphi}_n)$ is asymptotically normal under weak regularity conditions on the true data generating process. Crucially, this holds *regardless* of whether the likelihood is correctly specified: the required regularity conditions are effectively identical to those used to establish the asymptotic normality of the Frequentist quasi-maximum likelihood estimator. Hence, under mild conditions both the Bayesian posterior and Frequentist maximum likelihood estimator are asymptotically normal. Now, let J denote the information matrix, and let H denote the expected Hessian. When the information matrix equality $H = -J$ holds, the Bayesian posterior and Frequentist large-sample distributions agree: both have variance matrix J^{-1} . In this case, we appeal to Theorem 5 of [Kline and Tamer \(2016\)](#) to show that a $(1 - \delta)$ credible set for Θ is also an exact pointwise $(1 - \delta)$ Frequentist confidence set.²²

If the normal likelihood for the reduced form errors is correctly specified, then the information matrix equality holds. Correct specification, however, is not a necessary condition. Let \hat{s}_{ij} and \hat{s}_{lm} be the maximum likelihood estimators of two arbitrary elements s_{ij} and s_{jm} of the reduced form covariance matrix Σ . The necessary and sufficient condition for Bayesian posterior and Frequentist inference for φ to agree in our example is that the asymptotic covariance between \hat{s}_{ij} and \hat{s}_{lm} equals $(s_{ij}s_{jm} + s_{im}s_{jl})$. When this condition fails, the equivalence between credible sets and confidence intervals described in the preceding paragraph no longer holds. A solution to this problem is to avoid explicitly specifying a prior and likelihood and instead sample $\varphi^{(j)}$ from a multivariate normal distribution constructed to exactly match the Frequentist asymptotic distribution. This idea corresponds to the “pragmatic Bayesian” approach described by [Sims \(2010\)](#) and the “artificial ‘sandwich’ posterior” of [Müller \(2013\)](#). While we are in general supportive of this idea, we do not adopt it here for two reasons. First, implementing it in our examples would require us to rely on estimated fourth-order moments of the distribution of $(\varepsilon, \xi, \zeta)$, which are likely to be unreliable in practice. Second, our partial identification bounds rely crucially on the positive definiteness

²¹See [Hartigan \(1983\)](#) 4.4 for regularity conditions sufficient for Doob’s Theorem.

²²Formally, one must first verify an asymptotic independence property given in Assumption 5 of [Kline and Tamer \(2016\)](#). The examples considered in the present paper, however, fall under the case discussed in Remark 5 and Lemma 1 from [Kline and Tamer \(2016\)](#), so that one only requires the validity of both the usual Frequentist delta-method, and its Bayesian analogue.

of Σ , but drawing the half-vectorization of this matrix, $\text{vech}(\Sigma)$, from a multivariate normal distribution can produce draws that violate this restriction.

As in the classical measurement error case, we draw the reduced form covariance matrix from an Inverse-Wishart posterior when T^* is binary. Of course, the distribution of U cannot in fact be normal if any of the variables (y, T, z) is discrete. Nevertheless, the posterior for the reduced form parameters will still be asymptotically normal, centered at the maximum likelihood estimates. Provided that the aforementioned condition on the asymptotic covariance between \hat{s}_{ij} and \hat{s}_{lm} holds *approximately*, this asymptotically normal posterior will likewise approximate the Frequentist large-sample distribution. One could, in principle, write down a different likelihood for the binary T^* case. But this would require one to model the distribution of $T^*|\mathbf{x}$, an object over which applied researchers are typically agnostic when reporting OLS and IV results. For this reason, we prefer to treat the continuous and binary T^* cases within a common framework. Note, however, that the bounds for ψ from [Proposition 3.1](#) involve p . We suggest adopting an empirical Bayes approach and setting p equal to the sample analogue \hat{p} . This is irrelevant from a large-sample perspective, and amounts to a rounding error in applications. When the exogenous covariates \mathbf{x} include only a constant, p equals φ_T , so one could obtain posterior draws for this parameter directly from our normal-Jeffreys model. In the general case, however, it is less straightforward to obtain posterior draws for p . For one, the reduced form regression for T is not a generative model: it could imply conditional probabilities that are outside of $[0, 1]$. Addressing this difficulty would require one to either adopt a non-parametric approach or to impose parametric assumptions on the distribution of $T|\mathbf{x}$. Moreover, converting the conditional probability $\mathbb{P}(T|\mathbf{x})$ into the unconditional probability p requires integrating over the distribution of \mathbf{x} . The additional complications required to incorporate posterior uncertainty over p for the general \mathbf{x} seem excessive, particularly given that sampling uncertainty in p is of a smaller order than sampling uncertainty in Σ .